

IC549 U S. P10

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

"EXPRESS MAIL" MAILING LABEL

Number: EM 334 768 233 US Date of Deposit: October 29, 1977

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20590.
Typed or Printed Name: Teri Rinaldi

Signed:

**Docket
Number:**

P-65409/WHD/MTK
(Genentech Docket No.
PO1188)

Type a plus sign (+)
inside this box ~

+



INVENTOR(s)/APPLICANT(s)					
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)		
WOOD	WILLIAM	I.	San Mateo, California		
GURNEY	AUSTIN	L.	Belmont, California		
GODDARD	AUDREY		San Francisco, California		
TITLE OF THE INVENTION (280 characters max)					
NOVEL POLYPEPTIDES HAVING HOMOLOGY TO INTEGRIN AND NUCLEIC ACIDS ENCODING THEREFOR					
CORRESPONDENCE ADDRESS					
Walter H. Dreger, Registration No. 24, 190 FLEHR HOHBACH TEST ALBRITTON & HERBERT LLP, 4 Embarcadero Center, Suite 3400, San Francisco					
STATE	CA	ZIP CODE	94111	COUNTRY	US
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification	Number of Pages	39	<input type="checkbox"/>	Small Entity Statement
<input checked="" type="checkbox"/>	Drawings	Number of Sheets	35	<input type="checkbox"/>	Other (specify):
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the Provisional filing fees			PROVISIONAL FILING FEE AMOUNT (\$)	\$150.
<input checked="" type="checkbox"/>	The Commissioner is hereby authorized to charge any underpayment or overpayment to Deposit Account Number:			06-1300 (Order No. P-65409/WHDM/MTK)	

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

X

No.

Yes, the name of the U.S. Government Agency and the Government contract number are:

Respectfully submitted,

SIGNATURE:

By: J. P. Bourdelle Reg No. 35,081

Date _____

Oct. 29, 1997

For: TYPED or PRINTED NAME Walter H. Dreger

REGISTRATION NO.
(if appropriate)

24,190

Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take .2 hours to complete. Times will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Assistant Commissioner for Patents, Washington, DC 20231. **DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO:** Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

BEST AVAILABLE COPY



**NOVEL POLYPEPTIDES HAVING HOMOLOGY TO INTEGRIN
AND NUCLEIC ACIDS ENCODING THEREFOR**

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides having homology to integrin, designated herein as "PRO295" polypeptides.

BACKGROUND OF THE INVENTION

Extracellular proteins play an important role in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment.

Secreted proteins have various industrial applications, including pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.*, 93:7108-7113 (1996); U.S.

60063435-102997

Patent No. 5,536,637)].

The integrins comprise a supergene family of cell-surface glycoprotein receptors that promote cellular adhesion. Each cell has numerous receptors that define its cell adhesive capabilities. Integrins are involved in a wide variety of interaction between cells and other cells or matrix components. The integrins are of particular importance in regulating movement and function of immune system cells. The platelet IIb/IIIa integrin complex is of particular importance in regulating platelet aggregation. A member of the integrin family, integrin β -6, is expressed on epithelial cells and modulates epithelial inflammation. Another integrin, leucocyte-associated antigen-1 (LFA-1) is important in the adhesion of lymphocytes during an immune response. The integrins are expressed as heterodimers of non-covalently associated alpha and beta subunits. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being undertaken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification of a novel polypeptide which has homology to integrin.

SUMMARY OF THE INVENTION

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO295".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO295 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO295 polypeptide having amino acid residues 1 to 350 of Figure 2 (SEQ ID NO:3), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides a vector comprising DNA encoding a PRO295 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO295 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO295 and recovering PRO295 from the cell culture.

In another embodiment, the invention provides isolated PRO295 polypeptide. In particular, the invention provides isolated native sequence PRO295 polypeptide, which in one

embodiment, includes an amino acid sequence comprising residues 1 to 350 of Figure 2 (SEQ ID NO:3).

In another embodiment, the invention provides chimeric molecules comprising a PRO295 polypeptide fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO295 polypeptide fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to a PRO295 polypeptide. Optionally, the antibody is a monoclonal antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) containing the nucleotide sequence (SEQ ID NO:2) of a native sequence PRO295 cDNA (nucleotides 153-1202), wherein the nucleotide sequence (SEQ ID NO:1) is designated herein as "UNQ258" and/or "DNA38268". Also presented is the position of the initiator methionine residue. The putative transmembrane domain of the protein is encoded by nucleotides beginning at nucleotide 1586 in the figure.

Figure 2 shows the amino acid sequence (SEQ ID NO:3) derived from nucleotides 153-1202 of the nucleotide sequence shown in Figure 1.

Figures 3A-3B show an alignment of portions of the nucleotide sequences from a variety of expressed sequence tags as well as an intermediate consensus nucleotide sequence assembled therefrom designated as "DNA35814".

Figures 4A through 4AB show an alignment of portions of the nucleotide sequences from a variety of expressed sequence tags as well as an extended consensus nucleotide sequence assembled therefrom using repeated cycles of BLAST and phrap.

Figure 5 shows a BLAST sequence alignment analysis of portions of the PRO295 amino acid sequence derived from the DNA38268 molecule ("DNA38268") with portions of the integrin beta-6 subunit - cavia porcellus ("ITB6_CAVPO").

Figure 6 shows a BLAST sequence alignment analysis of portions of the PRO295 amino acid sequence derived from the DNA38268 molecule ("DNA38268") with portions of the lens fiber protein clfest4 precursor - gallus ("LFE4_CHICK").

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO295 polypeptide" and "PRO295" when used herein encompass native sequence PRO295 and PRO295 polypeptide variants (which are further defined herein). The PRO295 polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO295 polypeptide" comprises a polypeptide having the same amino acid sequence as a PRO295 polypeptide derived from nature. Such native sequence PRO295 polypeptide can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO295 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of a PRO295 polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of a PRO295 polypeptide. In one embodiment of the invention, the native sequence PRO295 polypeptide is a mature or full-length native sequence PRO295 polypeptide comprising amino acids 1 to 350 of Figure 2 (SEQ ID NO:3).

"Percent (%) amino acid sequence identity" with respect to the PRO295 amino acid sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO295 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

"Percent (%) nucleic acid sequence identity" with respect to the UNQ258 (DNA38268) sequence (SEQ ID NO:1) and nucleotides 153-1202 of the sequence shown in Figure 1 (SEQ ID NO:2) is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the UNQ258 (DNA38268) sequence (SEQ ID NO:1) or nucleotides 153-1202 of the sequence shown in Figure 1 (SEQ ID NO:2), respectively, after aligning the

sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO295 polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO295 polypeptide-encoding nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO295 polypeptide-encoding nucleic acid. An isolated PRO295 polypeptide-encoding nucleic acid molecule is other than in the former setting in which it is found in nature. Isolated PRO295 polypeptide-encoding nucleic acid molecules therefore are distinguished from the PRO295 polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated PRO295 polypeptide-encoding nucleic acid molecule includes PRO295 polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express PRO295 polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression

of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO295 polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO295 antibody compositions with polypeptidic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Active" or "activity" for the purposes herein refers to form(s) of PRO295 which retain the biologic and/or immunologic activities of native or naturally-occurring PRO295 polypeptide.

II. Compositions and Methods of the Invention

A. Full-length PRO295 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO295. In particular, Applicants have identified and isolated cDNA encoding a PRO295 polypeptide, as disclosed

in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO295 polypeptide have significant homology with the integrin protein as shown in Figures 5 and 6). Accordingly, it is presently believed that PRO295 polypeptide disclosed in the present application is a newly identified member of the integrin family and possesses cell adhesion typical of the integrin family.

B. Modifications of PRO295

Covalent modifications of PRO295 polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO295 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a PRO295 polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO295 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO295 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO295 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO295 polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO295 polypeptide.

50063435-102997

Addition of glycosylation sites to PRO295 polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO295 polypeptide (for O-linked glycosylation sites). The PRO295 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO295 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO295 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO295 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO295 comprises linking the PRO295 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

PRO295 polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a PRO295 polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a PRO295 polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the PRO295 polypeptide. The presence of such epitope-tagged forms of a PRO295 polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO295 polypeptide to be readily purified by affinity

purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a PRO295 polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

C. Preparation of PRO295

The description below relates primarily to production of PRO295 by culturing cells transformed or transfected with a vector containing at least nucleotides 153-1202 (SEQ ID NO:2) of the UNQ258 (DNA38268) nucleic acid (SEQ ID NO:1). It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO295 polypeptides. For instance, the PRO295 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of PRO295 polypeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce a full-length PRO295 polypeptide.

1. Isolation of DNA Encoding PRO295

DNA encoding a PRO295 polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the PRO295 mRNA and to express it at a detectable level. Accordingly, human PRO295-encoding DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO295-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to a PRO295 polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO295 is to use PCR methodology [Sambrook et al., *supra*; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., *supra*.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTAR, and INHERIT which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in

Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO295 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO_4 and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited

to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO295-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism.

Suitable host cells for the expression of glycosylated PRO295 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding the desired PRO295 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The desired PRO295 polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO295-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO295-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., *Nature*, 282:39 (1979);

Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO295-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO295 polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO295 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are

compatible with the host cell systems.

Transcription of a DNA encoding a PRO295 polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO295 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO295.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO295 polypeptides in recombinant vertebrate cell culture are described in Gething et al., *Nature*, 293:620-625 (1981); Mantei et al., *Nature*, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO295 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO295-encoding DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO295 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO295 polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO295 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO295 polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO295 polypeptide produced.

D. Uses for PRO295

Nucleotide sequences (or their complement) encoding PRO295 polypeptides have various applications in the art of molecular biology, including uses as hybridization probes,

in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO295-encoding nucleic acid will also be useful for the preparation of PRO295 polypeptides by the recombinant techniques described herein.

The full-length nucleotide sequence SEQ ID NO:1 or the full-length native sequence PRO295 (SEQ ID NO:2) nucleotide sequence, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO295 gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of PRO295 or PRO295 from other species) which have a desired sequence identity to the PRO295 sequence disclosed in Figure 1 (SEQ ID NO:1). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the UNQ258 (DNA38268) nucleotide sequence of SEQ ID NO:1 as shown in Figure 1 or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO295-encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO295 gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO295 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO295 sequences.

Nucleotide sequences encoding a PRO295 polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO295 polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO295 encode a protein which binds to another protein (example, where the PRO295 polypeptide functions as a receptor), the PRO295 polypeptide can be used in assays to identify the other proteins or molecules involved in the

binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO295 polypeptide can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO295 or a receptor for PRO295. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO295 polypeptide or any of its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO295 polypeptide can be used to clone genomic DNA encoding PRO295 in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO295. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO295 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO295 introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO295. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO295 can be used to construct a PRO295 "knock out" animal which has a defective or altered gene encoding PRO295 as a result of homologous recombination between the endogenous gene encoding PRO295 and altered genomic DNA encoding PRO295 introduced into an embryonic cell of the animal. For example, cDNA encoding PRO295 can be used to clone genomic DNA encoding PRO295 in accordance with established techniques. A portion of the genomic DNA encoding PRO295 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO295 polypeptide.

PRO295 polypeptides and portions thereof which have homology to integrin may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel integrins and integrin-like molecules may have relevance to a number of human disorders such as modulating the binding or activity of cells of the immune system. Thus, the identification of new integrins and integrin-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO 295.

E. Anti-PRO295 Antibodies

The present invention further provides anti-PRO295 polypeptide antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO295 antibodies of the present invention may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO295 polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO295 antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO295 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell

line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against a PRO295 polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

60063435-102997

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The anti-PRO295 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR

residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)].

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a PRO295 polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific

antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

F. Uses for anti-PRO295 Antibodies

The anti-PRO295 antibodies of the present invention have various utilities. For example, anti-PRO295 antibodies may be used in diagnostic assays for PRO295 polypeptides, e.g., detecting expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO295 antibodies also are useful for the affinity purification of PRO295 polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a PRO295 polypeptide are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO295 polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO295 polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO295 polypeptide from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1: Isolation of cDNA Clones Encoding Human PRO295

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington;

http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA35814 (see Figures 3A and 3B). In some cases, the consensus DNA sequence was extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended assembly sequence, is indicated as a number of second alignment figures, as shown in Figures 4A-AB.

Based on the DNA35814 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO295. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest by the *in vivo* cloning procedure using the probe oligonucleotide and one of the primer pairs.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCAGAGCGGAGATGCAGCGGCTTG-3'
(SEQ ID NO:4)

forward PCR primer (.f2) 5'-CCCAGCATGTACTGCCAG-3' (SEQ ID NO:5)

forward PCR primer (.f3) 5'-TTGGCAGCTTCATGGAGG-3' (SEQ ID NO:6)

forward PCR primer (.f4) 5'-CCTGGGCAAAAATGCAAC-3' (SEQ ID NO:7)

reverse PCR primer (.r1) 5'-CTCCAGCTCCTGGCGCACCTCCTC-3' (SEQ ID NO:8)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35814 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCTCTCAGCTACCGCGCAGGAGCGAGGCCACCCTCAATGAGATG-3'

(SEQ ID NO:10)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO295 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, **253**:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO295 [herein designated as UNQ258 (DNA38268)] (SEQ ID NO:1) and the derived protein sequence for PRO295.

The entire nucleotide sequence of UNQ258 (DNA38268) is shown in Figure 1 (SEQ ID NO:1). Clone UNQ258 (DNA38268) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 [Kozak et al., *supra*] and ending at the stop codon at nucleotide positions 1202-1205 (Figure 1). The predicted polypeptide precursor is 350 amino acids long (Figure 3). Clone UNQ258 (DNA38268) has been deposited with ATCC and is assigned ATCC deposit no. NEED DEPOSIT INFORMATION.

Analysis of the amino acid sequence of the full-length PRO295 polypeptide suggests that portions of it possess significant homology to the integrin proteins as shown in Figures 5 and 6, thereby indicating that PRO295 may be a novel integrin.

EXAMPLE 2: Use of PRO295-encoding DNA as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO295 as a hybridization probe.

DNA comprising the coding sequence of full-length PRO295 (as shown in Figure 1, SEQ ID NO:1) is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO295) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled UNQ258 (DNA38268)-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO295 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO295 Polypeptides in *E. coli*

This example illustrates the preparation of unglycosylated forms of PRO295 polypeptides by recombinant expression in *E. coli*.

The DNA sequence encoding PRO295 (SEQ ID NO:3) is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO295 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the

methods described in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO295 polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the polypeptide.

EXAMPLE 4: Expression of PRO295 Polypeptides in Mammalian Cells

This example illustrates preparation of glycosylated forms of PRO295 polypeptides by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO295-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO295-encoding DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO295

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO295 DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and

replaced with culture medium (alone) or culture medium containing 200 $\mu\text{Ci/ml}$ ^{35}S -cysteine and 200 $\mu\text{Ci/ml}$ ^{35}S -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO295 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO295-encoding DNA may be introduced into 293 cells transiently using the dextran sulfate method described by Sompayrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO295 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 $\mu\text{g/ml}$ bovine insulin and 0.1 $\mu\text{g/ml}$ bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO295 polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO295 polypeptide can be expressed in CHO cells. The pRK5-PRO295 vector can be transfected into CHO cells using known reagents such as CaPO_4 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S -methionine. After determining the presence of PRO295 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO295 polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO295 polypeptide may also be expressed in host CHO cells. The PRO295-encoding DNA may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO295-encoding DNA insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with

the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO295 polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

EXAMPLE 5: Expression of a PRO295 Polypeptide in Yeast

The following method describes recombinant expression of PRO295 polypeptides in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO295 polypeptide from the ADH2/GAPDH promoter. DNA encoding the PRO295 polypeptide of interest, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO295 polypeptide. For secretion, DNA encoding the PRO295 polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO295 polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO295 polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO295 polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 6: Expression of PRO295 Polypeptides in Baculovirus

The following method describes recombinant expression of PRO295 polypeptides in Baculovirus.

The PRO295-encoding DNA is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids

derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO295-encoding DNA or the desired portion of the PRO295-encoding DNA (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 to 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., *Baculovirus expression vectors: A laboratory Manual*, Oxford:Oxford University Press (1994).

Expressed poly-his tagged PRO295 polypeptide can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO295 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO295 polypeptide can

be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 7: Preparation of Antibodies that Bind PRO295 Polypeptides

This example illustrates the preparation of monoclonal antibodies which can specifically bind to PRO295 polypeptides.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO295 polypeptide, fusion proteins containing a PRO295 polypeptide, and cells expressing recombinant PRO295 polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO295 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO295 polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO295 polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO295 polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against a PRO295 polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO295 polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
-----------------	----------------------	---------------------

NEED DEPOSIT INFORMATION

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposit will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

60063435 . 1.02997

WHAT IS CLAIMED IS:

1. Isolated nucleic acid comprising DNA encoding a PRO295 polypeptide having amino acid residues 1 to 350 of Figure 2 (SEQ ID NO:3).
2. The nucleic acid of Claim 1 wherein said DNA comprises the nucleotide sequence of SEQ ID NO:1 or its complement.
3. The nucleic acid of Claim 1 wherein said DNA comprises nucleotides 153-1202 of the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2).
4. An isolated nucleic acid comprising the nucleotide sequence of the full-length coding sequence of clone UNQ258 (DNA38268) deposited under accession number ATCC
NEED DEPOSIT INFORMATION.
5. A vector comprising the nucleic acid of any one of Claim 1 or Claim 4.
6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.
7. A host cell comprising the vector of Claim 5.
8. The host cell of Claim 7 wherein said cell is a CHO cell.
9. The host cell of Claim 7 wherein said cell is an *E. coli*.
10. The host cell of Claim 7 wherein said cell is a yeast cell.
11. A process for producing a PRO295 polypeptide comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO295 polypeptide and recovering said PRO295 polypeptide from the cell culture.

12. Isolated native sequence PRO295 polypeptide comprising amino acid residues 1 to 350 of Figure 2 (SEQ ID NO:3).

13. A chimeric molecule comprising a PRO295 polypeptide fused to a heterologous amino acid sequence.

14. The chimeric molecule of Claim 13 wherein said heterologous amino acid sequence is an epitope tag sequence.

15. The chimeric molecule of Claim 13 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

16. An antibody which specifically binds to a PRO295 polypeptide.

17. The antibody of Claim 16 wherein said antibody is a monoclonal antibody.

60063435-102997

Abstract of the Disclosure

The present invention is directed to novel polypeptides having homology to the integrin protein and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

60063435 "102997

FIGURE 1A

CGCCGGCTCCGACCCCGGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCCGCAGC
 CCGCGGCTCCCGGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCA
 GTCGGGCGGGCTGCGGGCGCAGAGCGGAG
 ><MET {trans=1-s, dir=f, res=1}
 ATGACGGCTTGGGCCACCTGTGTGCTGCTGTGGCGGGCGGTCCCCACGGCC
 CCGCGGCTCCGACCGCGACCTCGGCTCAGTCAAGCCCGGCCCGCTCTCAGCTAC
 CCGCAGGAGGCGCACCTCAATGAGATGTTCCGCGAGGTTGAGGAACCTGATGGAGGAC
 ACGCAGCAAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCGAGAAGCTGCTGTAA
 GCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACA
 GACACGAAGGTTGGAATAATACCATCTGTGACCGAGAAATTCACAAGATAACCAAC
 AACACAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGAGACGAAGAA
 GGAGAGAGGCCACGAGTGATCATCGACGAGGACTGTGGGCCCGCAGCATGTACTGCCAG
 TTTGCCAGCTTCCAGTACACCTGCCAGCATCGCGGGCCAGAGGATGCTTGCACCCGG
 GACAGTGTGTGTGGAGACCACTGTGTGTCTGGGTCACTGCACCAAAATGGCCACC
 AGGGCAGCAATGGGACCATCTGTGACAAACAGAGGGACTGCCAGCCGGGGCTGTGTGT
 GCCTTCCAGAGAGGCTGTGTCTCCCTGTGTGCACACCCCTGCCGTGGAGGCGAGCTT
 TGCCATGACCCCGCCAGCGGCTTCTGACCTCATCACCTGGGAGCTAGAGCTGATGGA
 GCCTTGGACCGATGCCCTTGTGCCAGTGGCTCTCTGCCAGCCCAAGATGGGAGATCCTGTG
 GTGTATGTGTCAAGCCGACCTTCTGTGGGAGCCGTGACCAAGATGGGAGATCCTGTG
 CCCAGAGAGTCCCAGATGATGAAAGTTGGCAGCTTATGAGAGAGATGGCGCTGGGGAGCCTGCCGCTGCC
 CTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGAGCCTGCCGCTGCC
 GCCGTGCACTGCTGGAGGGGAAGAGATTTAGATCTGGAACAGGCTGTGGGTAGATGTG
 CAATAGAAATAGCTAAATTTATTTCCCGAGGTGTGTCTTAGCGTGGCTGACCAGGCT
 TCTTCTACATCTTCTCCAGTAAAGTTTCCCTCTGGCTTTCACAGTCTGGTGTGGAGATC
 CATTTGTTAGCTCCCGAGGCTGTTCTCCAGGCTTTCACAGTCTGGTGTGGGAGATC
 AGCAGGGTTAACTGCAGGAGCAGTTTGGCCACCCCTGTCCAGATTATTGGCTGTTC
 CTCTACCAAGTTGGCAGACAGCGTTTGTCTACATGGCTTTGATAATTGTTGAGGGAG
 GAGATGGAAACAATGTGGAGTCTCCCTGTGATTGGTTTGGGAAATGTGGAGAAGATG
 CCCTGCTTTGCAACAATCAACCTGGCAAAATGCAACAATGAATTTTCCACGAGTCT
 TTCCATGGGCATAGGTAAGCTGTGCTTCAGCTGTTCAGATGAATGTTGTTTACCCC
 TGCATTACATGTGTTATTCATCCAGCAGTGTGCTCAGCTCTACCTGTGTCAGGGC
 AGCATTTTCATATCCAGATCAATTCCTCTCTCAGCAGCCTGGGAGGGGTCATTG
 TTTCTCTGTCATCAGGGATCTCAGAGGCTCAGAGACTGAAGCTGCTTGCCTCAAGTCA
 CACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCTCT
 CCACCTACCCACACACAGCCTTGGTGCCCAAAAGTGTCTCCCCAAAGGAAGGAAATGG

FIGURE 1B

GATTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACTAATTCTCATCCCTCTA
 AAGTAAACTACTGTAGGAACAGCAGTGTCTCAGTGTGGGCAGCCGTCTCTCTAA
 TGAAGACAAATGATATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACTTTGAAAGGT
 ATATGACTGAGCGTAGCATACAGGTTAACTTCAGAAACAGTACTTAGGTAATTGTAGGG
 CGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAACTGAAGACAATATCAACC
 ACGTGGAGAAATCAAAACCGAGCAGGGCTGTGTGAAACATGGTTGTAATATGCCACTGCCG
 AACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCTATGGACTGTTGCCAC
 CATGTAATTCATCCAGAGTCTTAAAGTTTAAAGTTGCACATGATGTATATAAGCATGCTTT
 CTTTGAGTTTAAATATGTATAAACATAAAGTTGCATTTAGAAATCAAGCATAAATCACT
 TCAACTGCAAAAAAATAAAAAAATAAAAAA

466207" SETE9009

FIGURE 2

MQRLGATLLCLLLAAVPTAPAPAPTATSA PVKPGPALSYPOEEATLNEMFREVEELMED
TQHKLRSAVEEMEAEAAAAKASSEVNLANLP SYHNETNTDTKVGNNTIHVHREIHKITN
NQTGQMVFS ETVITSVGDEEGRRSHECI IDEDCGSPMYCQFASFYTCQPCRQRM LCTR
DSECCGDQLCVWGHCTKMATRGNGTI CDNRDCQPGLC CAFORGLFPVCTPLPVEGEL
CHDPASRLLDLITWELEPDGALDRCP C ASGLLCQPHSHSLVYVCKPTFVGS RDQDGEILL
PREVPDEYEVGSFMEEVRQ ELEDLERSLTEEMALGE PAAAAAALLGEEI

FIGURE 3A

2054176	1	CTGTGGGAGACGAAGGAGGAGGAGCCACGAGTGTCATCATCGACGAG
1756022	1	TGTGGGAGACGAAGGAGGAGGAGCCACGAGTGTCATCATCGACGAG
980963	1	GACGAAGAAGGAGGAGGAGGAGCCACGAGTGTCATCATCGACGAG
W61032	1	GCCACGAGTGTCATCATCGACGAG
<DNA35814:	1	CTGTGGGAGACGAAGGAGGAGGAGGAGCCACGAGTGTCATCATCGACGAG
2054176	51	GACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTG
1756022	50	GACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTG
980963	43	GACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTG
W61032	24	GACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTG
853079	1	AGTACACCTG
<DNA35814:	51	GACTGTGGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTG
2054176	101	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
1756022	100	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
980963	93	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
W61032	74	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
853079	11	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
<DNA35814:	101	CCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTNAGTGCT
2054176	151	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
1756022	150	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
980963	143	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
W61032	124	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
853079	61	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
<DNA35814:	151	GTGGAGACCAGCTGTGTGCTGTTGGGTCACTGCACCAAAATGGCCACCAGG
2054176	201	GGCAGCAATGGGACCACCTCTGTGACAACCAAGAGGGGACTGCCA
1756022	200	GG
980963	193	GGCAGCAATGGGACCACCTCTGTGACAACCAAGAGGGGACTGCCAGCCGGGGCT
W61032	174	GGCAGCAATGGGACCACCTCTGTGACAACCAAGAGGGGACTGCCAGCCGGGGCT
853079	111	GGCAGCAATGGGACCACCTCTGTGACAACCAAGAGGGGACTGCCAGCCGGGGCT
<DNA35814:	201	GGCAGCAATGGGACCACCTCTGTGACAACCAAGAGGGGACTGCCAGCCGGGGCT

FIGURE 3B

980963	243	GTGCTGTGCCCTTCCAG
W61032	224	GTGCTGTGCCCTTCCAGAGAGGCCCTGCTGTTCCCTGTGTGCACACCCCTGC
853079	161	GTGCTGTGCCCTTCCAGAGAGGCCCTGCTGTTCCCTGTGTGCACACCCCTGC
<DNA35814:	251	GTGCTGTGCCCTTCCAGAGAGGCCCTGCTGTTCCCTGTGTGCACACCCCTGC
W61032	274	CCGTGGAGGN-GAGCTTTGCCATG
853079	211	CCGTGGAGGGCGAGCTTTGCCATG
<DNA35814:	301	CCGTGGAGGGCGAGCTTTGCCATG

[illegible]

FIGURE 4E

AA226979	331	GAAATAATACCATCCATGTGCACCGAGAA
1452718	96	GAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
2845278	92	GAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
923337	38	GAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
T08793	4	GAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
T30923	1	TACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
981001	1	ATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
261347	1	CCAGAGGAGAGAAATTCACAAGATAACCAACAAC
T31076	1	GAGAAATTCACAAGATAACCAACAAC
<consen02>		+++++.....+++++.
1452718	353	GAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAAC
2845278	146	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
923337	142	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
T08793	88	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
T30923	54	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
981001	44	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
261347	40	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
T31076	33	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
1756022	27	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
1549807	1	ATCACATNTGTGGGAGA
DNA35814.init	1	CACATCTGTNGGANA
2054176	1	CTGTGGGAGA
980963	1	CTGTGGGAGA
<consen02>		+++++.+++++.
1452718	403	CAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGA
2845278	196	CGAAGAAGGCAGAGAGGCCACGAGTGCATCATCGACGAGGACTGT-GGG
923337	192	CGAAGAAGGCAGAGAGGCCACGAGTGCATCATCGACGAGGACTGT-GGG
T08793	138	CGAAGAAGGCAGAGAGGCCACGAGTGCATCATCGACGAGGACTGT-GGG
T30923	104	CGAAGAAGGCAGAGAGGCCACGAGTGCATCATCGACGAGGACTGT-GGG
981001	94	CGAAGAAGGCAGAGAGGCCACGAGTGCATCATCGACGAGGACTNT-GGG
	90	CGGAGAAGGCAGAGAGGAGNACGGGTNCAATATTGGCGGGGCTTTTGGG

FIGURE 4J

2658126	131	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT	131	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT
1478240	131	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT	131	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT
2524907	119	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT	119	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT
925806	94	GGCGA-GNTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATNACCN	94	GGCGA-GNTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATNACCN
W45085	84	CG-GACGCTTTGC-ATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT	84	CG-GACGCTTTGC-ATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT
1865280	66	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT	66	GGCGA-GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT
		...+ +.++++.+.+++++++.+++++.+++++.+++++.++++.		...+ +.++++.+.+++++++.+++++.+++++.+++++.++++.
<consen02>	751	GGCGA GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCT		
W61032	330	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
2482612	204	GGG-AGCTA		
R14945	200	GGGAGCTAGAGCCTGATGGAGCCTTGGGACCGATGCCCTTGTGCCAGTG		
1349102	185	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
2656314	180	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
2658126	180	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
1478240	180	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
2524907	168	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
925806	143	NGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGANGCCNTTGTGCCAGTG		
W45085	132	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
1865280	115	GGG-AGCTAGAGCCTGATGGAGCCTTGG-ACCGATGCCCTTGTGCCAGTG		
865729	1	CGATGCCCTTGTGCCAGTA		
2760910	1	CCCTTGTGCCAGTA		
2496946	1	TTGTGCCAGTN		
		.++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++ ++++++.		
<consen02>	800	GGG AGCTAGAGCCTGATGGAGCCTTGG ACCGATGCCCTTGTGCCAGTG		
W61032	378	GCCTCCTCT-GCCAGCCCCACAGCCACAGCCTGG-TGTATGT-GTGCAA-		
R14945	250	GCCTCCTCTTGCCAGCCCCACAGCCACAGCCTGGTGTATGTGTGCAAA		
1349102	233	GCCTCCTCT-GCCAGCCCCACAGCCACAC		
2656314	228	GCCT		
2658126	228	GCCTCCTCT-GCCAGCCCC		
1478240	228	GCCTCCT		
2524907	216	GCCTCCTCT-GCCAGCNCACACA		
925806	191	GCCTACTAN-GNCAGCCCCACAGCCACAGCCTGG-TNTATGT-GTGCAA-		

FIGURE 4K

W45085	180	GC-TCCTCT-GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
1865280	163	GCCTCCTCT-GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
865729	20	GNCTNCTNT-GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
2760910	15	GCCTCCTCT-GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
2496946	12	GCCTCCTCT-GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
2223827	1	GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
2226521	1	GCCAGCCCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
1466083	1	CCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
927569	1	CCCACAGCCACAGCCCTGG-TGTATGT-GTGCAA-
<consen02>	848	GCCTCCTCT GCCAGCCCCACAGCCACAGCCCTGG TGTATGT GTGCAA
W61032	424	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATC-TGCTGCCC
R14945	300	GCCGACCTTCGTNGGGGAACCGTGACCAAGATGGGGAGATTCCTT
925806	237	GCCGACCTTAGT-GGGAGCCCGTGACCAAGATGGGGAGATN-TG-TGCCC
W45085	225	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
1865280	209	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGG
865729	66	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
2760910	61	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
2496946	58	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
2223827	38	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
2226521	38	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
1466083	32	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
927569	32	GCCGACCTTCGT-GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
354028	1	CCGTGACCAAGATGGGGAGATCCTGCTGCCC
2277563	1	CCGTGACCAAGATGGGGAGATCCTGCTGCCC
1289239	1	CCGTGACCAAGATGGGGAGATCCTGCTGCCC
AA031480.RC	1	CCGTGACCAAGATGGGGAGATCCTGCTGCCC
<consen02>	894	GCCGACCTTCGT GGGAGCCCGTGACCAAGATGGGGAGATCCTGCTGCCC
W61032	472	A-GAGAGGTCCC-GAT-GAGTATGAAGTTGGAA--CTTCATGGAGGAGGT
925806	284	A-GAGAGGTTC-GAT-NAGTATGA
W45085	274	AANAAGGTCCCGATTGAGTATGAAGTTGGCAAGCTTCATGGAAGGAAN

FIGURE 4N

2760910	256	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAA
2496946	253	TTAGATCTGGACCAAG-CTGTGGGTAGATGTGCAATAGAAA-CTAATTT
2223827	233	TTAGATCTGGACCAAGGCTG
2226521	233	TTAGATCTGGACCAAGGCTG
927569	227	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCA
354028	178	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2277563	178	TTAGATCTGGACCAAGGCTGTGGGTAAATGTGCAATAGAAAATAGCTAATTT
1289239	177	TTAGATCTGGACCAAGGCTGTGGGTAGATGTNCAATAGAAAATAGCTAATTT
AA031480.RC	155	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
R52311	125	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
T33818	115	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
3087465	100	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2525451	90	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
C05296	84	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
C05172	81	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2189372	80	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
HSC0CH081	76	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
1369095	53	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2778468	30	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
T18925	28	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2885667	1	ATCTGGACCAAGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
H07079	1	TCTGGACC-GGNTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
R15488	1	GCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
T80210	1	GTCAANTAGAAAATAGCTAATTT
H18569	1	TCAATTAGAAAATAGCTAATTT
<consen02>	1089	TTAGATCTGGACCAAGGCTGTGGGTAGATGTGCAATAGAAAATAGCTAATTT
2496946	301	ATTTCCCAAGGTGTGCTTTA
354028	228	ATTTCCCAAGGTGTGCTTTAGGCGTGGGTGACCAAGG-TTCTTCCTAC
2277563	228	ATTTCCCAAGGTGTGCTTTAGGCGTGGGTGACCAAGG
1289239	227	ATTTCCCAAGGTGTGCTTTAGGCGTGGGTGACCAAGGCT
AA031480.RC	205	ATTTCCCAAAAAA
R52311	175	ATTTCCCAAGGNTGTGCTTTAGGCGTGGGTGACCAAGGCTTCTTCCNAC

FIGURE 40

T33818	165	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTNCCTAC
3087465	150	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
2525451	140	AAAAA
C05296	134	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
C05172	131	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
2189372	130	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTNCCTAC
HSC0CH081	126	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
1369095	103	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
2778468	80	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
T18925	78	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
2885667	47	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
H07079	45	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
R15488	36	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
T80210	23	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
H18569	22	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
045844	1	CCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
1704942	1	GTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
877050	1	CTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
865207	1	GGCGTGGGCTGACCAAGGCTTCTTCCCTAC
833987	1	GGCGTGGGCTGACCAAGGCTTCTTCCCTAC
2500233	1	GGCGTGGGCTGACCAAGGCTTCTTCCCTAC
N86855	1	GCTGACCAAGGCTTCTTCCCTAC
1405577	1	ACCAAGGCTTCTTCCCTAC
2056655	1	CCAGGCTTCTTCCCTAC
945395	1	CCAGGCTTGTTCCTAC
354502	1	AGGCTTCTTCCCTAC
<consen02>	1139	ATTTCCCCAGGTGTGCTTTAGGCGTGGGCTGACCAAGGCTTCTTCCCTAC
354028	277	ANCTTCTTCCCAGTAAGTTTCCCCTCTGGCTT-NACANCATGAGGTGTTG
R52311	225	ATCTTCTTCCCAGTAAGTTTCCCCTCTGGCTT-GACAGCATGAGGTGTTN
T33818	215	ATCTTCTTCCCAGTAAGTTTCCCCTCTGGCTT-GACAGCATGAGGTGTTG
3087465	200	ATCTTCTTCCCAGTAAGTTTCCCCTCTGGCTT-GACAGCATGAGGTGTTG
C05296	184	ATCTTCTTCCCAGTAAGTTTCCCCTCT

FIGURE 40

354028	326	TGCAATTGTTCAACTTNC	CCCCCANGNTTGT	TTTCCANGGTT	CANANTCTT
R52311	274	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
T33818	264	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGG	
3087465	249	TGCATT-GTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTTC	
C05172	230	TGCATTTTGTTCAGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
2189372	229	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
HSC0CH081	225	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
1369095	202	TGCATTTGTTACG			
2778468	179	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
2885667	146	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
H07079	144	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
R15488	135	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
T80210	123	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
H18569	121	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
045844	94	TGCATTTTTCAGNT--	NCCCCAGGCT--	GTTCTCCAGGCTT	NANAGTCT-
1704942	90	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
877050	83	TGCATTTGTTACGCT--	CCCCCAGGNT--	GTTCTCCAGGCTT	CACAGTCT-
865207	78	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
833987	77	TGCATTTGTTNAGCT--	CCCCCAGGCT--	GTTCTNCAGGCTT	CACAGTCT-
2500233	73	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
N86855	71	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
1405577	67	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
2056655	66	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
945395	66	TGCATTTGTTACGNT--	NCCCCAGGCT--	GTTNTNCAGGCTT	CACAGTNT-
354502	64	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
681862	48	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTNTCCAGGNTT	CACAGTCT-
2547081	32	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
381195	29	TGCATTTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGCTT	CACAGTCT-
947376	2	TGCATTTNTTCAGNT--	CCCCNAGGCT--	GTTCTCCAGGCTT	NACAGTCT-
1351726	1	CGGCTCGAGCTCGAG--	CCGATTGGC-	TCGAGCCAGGCTT	CACAGTCT-
624084	1	ATTGTTACGCT--	CCCCCAGGCT--	GTTCTCCAGGNTT	CACAGTCT-
622154	1	ATTGTTACGCT--	CCCCCAGGNT--	GTTCTNCAGGNTT	CACAGTCT-
1727802	1			CTCCAGGCTT	CACAGTCT-
HSC3MH091	1			CTCCAGGCTT	CACAGTCT-

FIGURE 4U

354502	204	AGCCGTTTGTCTACATGG-CTTTGATAA-TT
681862	188	AGCCGTTTGTCTACATGG-NTTTGATAA-TTGT-TTG
2547081	172	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
381195	169	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
947376	142	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
1351726	141	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
624084	138	AGCCGTTTGTCTACATGG-NTTTGATAA-TTGT-TTGAG-GGG-AGG-A
622154	138	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAN-GGG-AGG-A
1727802	113	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGNG-GGG-AGG-A
HSC3MH091	113	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGN-TTNG-GGG-ANG-A
2156323	93	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGN-TTNG-GGG-AGG-A
463503	87	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
T78829	80	AGC-GTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
HSC37E101	76	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
465657	72	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
N36746	56	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
T65270	49	AGC-GTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
1503017	46	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAG-GGG-AGG-A
1370618	38	AGCCGTTTGTCTACATGG-NTTTGATAA-TTGT-TTGNG-GGG-AGG-A
712299	21	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGNGCCGG-AGG-A
2899383	10	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAN-GGG-AGG-A
2666235	11	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAN-GGG-AGG-A
2846187	4	AGCCGTTTGTCTACATGG-CTTTGATAA-TTGT-TTGAN-GGG-AGG-A
1358133	1	GG-A
1349616	1	GG-A
<consen02>		++..++++.++...+++++++ .+++..++++ ++.. .+.... ++. +.+ +
	1378	AGCCGTTTGTCTACATGG CTTTGATAA TTGT TTGAG GGG AGG A

FIGURE 4V

H07079	332	GATGGGAACAATGTGGAGTCTCCCTCT-GATTGGTTTGGGGGAAATG
R15488	315	GATGG-AA-CAATGTGGAGT
T80210	307	GATGG-AAACAATGTGGAGTCTCCCTCT-GATTGGTTTGGGG-AAATG
H18569	305	GATGG-AAACAATGTGGAGTCTCCCTCT-GATTGGTTTGGGG-AAATG
865207	262	G
833987	260	GATGG-AAAC
N86855	254	GATGG-AAACAATGTGG-AGTCTCCCTCTTGAT-GGTTT-GGGG-AAAGT
2056655	250	GATGG-AAACAATGTGG-AGTCTCCCT
945395	250	GATGG-AAACAATGTGG-AGTCTCCCTGT-GATTGGT
2547081	216	GATGG-AAACAATGTGG-AGTCTCCCTCT-G
381195	213	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
947376	186	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATN
1351726	185	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
624084	182	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
622154	182	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
1727802	157	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
HSC3MH091	157	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2156323	138	GATGG-AAANNANGGAGCNCNGGCTT-TCCCGTTGGG
463503	131	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
T78829	123	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
HSC37E101	120	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
465657	116	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
N36746	100	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
T65270	92	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
1503017	90	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
1370618	82	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
712299	65	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2899383	55	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2666235	55	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2846187	48	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
1358133	7	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
1349616	4	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2305359	1	GATGG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG
2149106	1	GG-AAACAATGTGG-AGTCTCCCTCT-GATTGGTTTGGGG-AAATG

FIGURE 4X

2305359	47	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
2149106	44	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
1984534	44	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
353611	35	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
052714	24	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
AA041443	24	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
AA090223	8	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
2152542	8	TGG-AGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
N26110.RC	1	GAAAG-AGTCCCTGCTTT-GCAAACATCAAACCTGGCAAAAATN
857436	1	GGAGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGGCAAAAATG
1377826	1	GGAGAAG-AGTGCCCTGCTTT-GCAAACATCAA-CCTGG-CAAAAATG
		.++...+.+ ++++.++++.++++ ..++++.+++++ .+.+.+.+.+.+.+
<consen02>	1469	TGGGAGAAG AGTGCCCTGCTTT GCAAACATCAA CCTGGGCAAAAATG
H07079	430	CAACAA
T80210	402	CAACAA-T-GAATTTTCC-ACGNAGTT
H18569	400	CAACAA-TTGAATTTTCC-NCGCANTTTCCTCC-CTTGGGCA-TAA
947376	277	CAACAA-T-GTNTTTT-CC-GNGGTGTT-N-TTNCC-NG-GG
HSC3MH091	248	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
463503	222	CAACAA-T-GNT
T78829	214	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
HSC37E101	211	CAACAA-T-GAATTTT-CC-ACGCA-TT-C-TTTC-AT-GGGCA-TA-
465657	207	CAACAA-T-GAAT
N36746	191	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
1503017	181	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
1370618	173	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
712299	156	CAACAA-T-GAATTTT-NC-ACGNAGTT-C-TTTC-AT-GGGCA-TA-
2899383	146	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
2666235	139	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
2846187	98	CAACAA-T-GAATTTT-CC-ACGCANTT-C-TTTC-AT-GGGCA-TA-
1358133	95	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
1349616	92	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-
2305359	89	CAACAA-T-GAATTTT-CC-ACGCAGTT-C-TTTC-AT-GGGCA-TA-

FIGURE 4X

1984534	89	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
353611	80	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
052714	69	CAACAAA-T-GANTTTC-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
AA041443	69	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
AA090223	53	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
2152542	53	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
N26110.RC	45	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
857436	46	CAACAAA-T-GAATTTT-CC-ACGNAGTT-C-TTTCC-AT-GGGCA-TA-
1377826	45	CAACAAA-T-GAATTTT-CC-ACGCAGTT-C-TTTCC-AT-GGGCA-TA-
		++++.++ + +.+++ . + .+.+++ . ++.++ . . +++ . ++
<consen02>	1516	CAACAAA T GAATTTT CC ACGCAGTT C TTTCC AT GGGCA TA
H18569	446	GGTAAGCTGTGCCCTTCAAGCTGTTTGGCCAGAGNAATGTTCTGGTT
HSC3MH091	288	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
T78829	254	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
HSC37E101	250	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
N36746	231	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
1503017	221	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
1370618	213	GGTAAGCTGT-GCC-TTCA-GC
712299	196	GGTAAGCTGT-GCC-TTNA-GCTGTTTTCAGATGAAAGG
2899383	186	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
2666235	186	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
2846187	179	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
1358133	139	NGTAACTGT-GNC-NTCA-NCTGTTGCAGATGAAATGTTCTGTTACCCC
1349616	135	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
2305359	132	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
2149106	129	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
1984534	129	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
353611	120	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
052714	109	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
AA041443	109	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
AA090223	93	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
2152542	93	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC
N26110.RC	89	GGTAAGCTGT-GCC-TTCA-GCTGTTGCAGATGAAATGTTCTGTTACCCC

FIGURE 4AA

T78829	350	TGTGCCAGGGCAGCCATTTTCATATCCCAAGATCAATTCCTC-TTTC
N36746	327	TGTGNCAGGG-CAGC-ATTTTCATATCC-AAGATCAATTCCTCCTCT-C
1358133	235	TGTGCCAGGG-CAGC-ATTTTCATAT
1349616	231	TGTGCCAGGG-CAGC-ATTTTCCTT
2305359	228	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGAT
2149106	225	TG
1984534	225	TGTGCCAGGG-CAGC-ATTTTCATATCC-NAGATCAATNCC
353611	216	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGATCAT
052714	205	TNTNCCAGGG-CAGC-ATTTTNATATCC-AAGATC
AA041443	205	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGATCAATTCCTC-TCT-C
AA090223	189	TGTGCCAGG--CAGC-ATTTTAATATCC-AGGATCAATTCCTC-TCT-G
2152542	189	TGTGCCAGGG-CNNN-ATTTTCATATCC-AAGATCAATTCCTC-TCT-C
N26110.RC	186	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGATCAATTCCTC-TCT-C
857436	182	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGATCAATTCCTC-TNT-C
1377826	181	TGTGCCAGGG-CAGC-ATTTTCATATCC-AAGATCAATTCCTC-TCT-C
		+..+..+++++ +... ++++++..+++++ ..+++++.+.+++++.+.+
<consen02>	1652	TGTGCCAGGG CAGC ATTTTCATATCC AAGATCAATTCCTC TCT C
T78829	399	AGGCACAGCCNGGGGAGGGGGTCTTG-TTCTTCCTCG-TCCATCAGG
N36746	373	AG-CACAGCCTGGGGAGGGGGTCTTGTTCT-CCTCG-GCCACACAGG
AA041443	250	AG-CACAGCCTGGGG-AGGGGG-TCATTG-TTCT-CCTCGNTCCATCAGG
AA090223	233	AG-CACA
2152542	234	AG-CACAGCCTGGGG-AGGG
N26110.RC	231	AG-CACAGCCTGGGG-AGGGGG-TCATTG-TTCT-CCTCG-TCCATCAGG
857436	227	AG-CACAGC-TGGGG-G
1377826	226	AG-CACAGCCTGGG
		++ ++++++.++++.+++++ ++++++ +++ +++++.++++.+
<consen02>	1697	AG CACAGCCTGGGG AGGGGG TCATTG TTCT CCTCG TCCATCAGG

FIGURE 5

ITB6_CAVPO Integrin beta-6 subunit - cavia porcellus, fragment (577 aa)
 Score = 113 (39.8 bits), Expect = 0.038, P = 0.037
 Identities = 38/119 (31%), Positives = 53/119 (44%), at 612,393, Frame = +3

DNA38268	612	GPSMYCQFASFQYTCQPCRGQRMCLCTRDSEC-CGDQLCVWG-----HCTKMATR--GSN
ITB6_CAVPO	393	GP--YCQCDNF--SCVRHKG--LLCGDNGDCECGECVCRSGWTGEYCNCITSTDTTCISED
DNA38268	765	GTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPPASRLDLITWELEPDG----
ITB6_CAVPO	447	GTLCSGRGDCVCGKCVCTNPGASGPTCERCPT-----CSDPCNSKRSCIECHLSADGQPGE
DNA38268	933	-ALDRCPCA
ITB6_CAVPO	503	ECVDKCKLA

FIGURE 6

LFE4_CHICK Unknown lens fiber protein clfest4 precursor - gallus g (350 aa)
 Score = 1003 (353.1 bits), Expect = 2.1e-100, P = 2.1e-100
 Identities = 180/293 (61%), Positives = 222/293 (75%), at 285,37, Frame = +3

DNA38268	285	ATLNEMFREVEELMEDTQHKLRSAVEEMEAEEAAKASSEVNLNLPSPSYHNETNTDTKV
LFE4_CHICK	37	ASLGEMLRVEALMEDTQHKLRNAVQEMEAEEEGAKKLSEVNFENLPPTYHNESNTETRI
DNA38268	465	GNNTIHVHREIHKITTNNQTGOMVFSETVITSVGDEEGRRSHECIIDEDCGPSPMYCQFASF
LFE4_CHICK	97	GNKTVQTHQEIDKVTDNRTGSTIFSETIITSIKGGENKRNHECIIDEDCETGKYCQFSTF
DNA38268	645	QYTCQPCRGQRMCLTRDSECCGDLQCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQR
LFE4_CHICK	157	EYKCQPCKTQHTHCSRDECCGDLQCVWGECKRATSRGNGTICENQHDGNPGTCCAFQK
DNA38268	825	GLLPVCTPLPVEGELCHDPASRLLDLITWELEPDGALDRCPGASGLLCQPHS-HSLVYV
LFE4_CHICK	217	ELLFPVCTPLPEEGEPCHDPNRLNLITWELEPDGVLERCPGASGLICQPQSSHSTTSV
DNA38268	1002	CKPTFVGSRDQDGE-----ILLPREVPDEYEVGSFMEEVRQEELEDLE
LFE4_CHICK	277	CELSNETRKNKEKEDPLNMDEMPFISLIPRDLSDYEESVIOEVRKELESLE

H/PKOV 1

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

"EXPRESS MAIL" MAILING LABEL

Number: EM 334 768 033 US Date of Deposit: October 29, 1997

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231 on October 29, 1997.

Typed or Printed Name: Lisa Lee

Signed: *Lisa Lee*



10/29/97

Docket Number:	P-65401/WH/D/DAV	Type a plus sign (+) inside this box -	+
----------------	------------------	--	---

INVENTOR(s)/APPLICANT(s)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Wood	William	I.	San Mateo, California
Goddard	Audrey		San Francisco, California
Gurney	Austin	L.	Belmont, California

TITLE OF THE INVENTION (280 characters max)			
NOVEL POLYPEPTIDES HAVING HOMOLOGY TO CD44 ANTIGEN AND NUCLEIC ACIDS ENCODING THE SAME			

CORRESPONDENCE ADDRESS					
FLEHR HOHBACH TEST ALBRITTON & HERBERT LLP, 4 Embarcadero Center, Suite 3400, San Francisco					
STATE	CA	ZIP CODE	94111	COUNTRY	US

ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification	Number of Pages	38	<input type="checkbox"/>	Small Entity Statement
<input checked="" type="checkbox"/>	Drawings	Number of Sheets	14	<input type="checkbox"/>	Other (specify):

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)			
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the Provisional filing fees	PROVISIONAL FILING FEE AMOUNT (\$)	\$150
<input checked="" type="checkbox"/>	The Commissioner is hereby authorized to charge any deficiencies in filing fees or credit any overpayments to Deposit Account Number: 06-1300 (Order No. P-65401/WH/D/DAV)		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

<input checked="" type="checkbox"/>	No.
<input type="checkbox"/>	Yes, the name of the U.S. Government Agency and the Government contract number are: _____

Respectfully submitted,

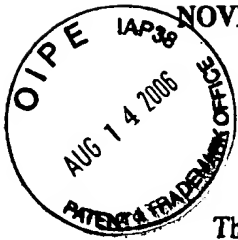
SIGNATURE: *Walter H. Dreger, Esq.*

TYPED or PRINTED NAME: Walter H. Dreger, Esq.

Date: October 29, 1997

REGISTRATION NO. 24,190
(if appropriate)

50064215-102997



NOVEL POLYPEPTIDES HAVING HOMOLOGY TO CD44 ANTIGEN AND
NUCLEIC ACIDS ENCODING THE SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA
5 and to the recombinant production of novel polypeptides having homology to CD44 antigen,
designated herein as "PRO263" polypeptides.

BACKGROUND OF THE INVENTION

CD44 is a cell surface adhesion molecule involved in cell-cell and cell-matrix
10 interactions. Hyaluronic acid, a component of the extracellular matrix is a major ligand.
Other ligands include collagen, fibronectin, laminin, chondroitin sulfate, mucosal addressin,
serglycin and osteopontin. CD44 is also important in regulating cell traffic, lymph node
homing, transmission of growth signals, and presentation of chemokines and growth factors
to traveling cells. CD44 surface proteins are associated with metastatic tumors and CD44 has
15 been used as a marker for HIV infection. Certain splice variants are associated with metastasis
and poor prognosis of cancer patients. Therefore, molecules having homology with CD44 are
of particular interest, as their homology indicates that they may have functions related to those
functions of CD44. CD44 is further described in U.S. Patent Nos. 5,506,119, 5,504,194 and
5,108,904; Gerberick, et al., Toxicol. Appl. Pharmacol., 146(1):1 (1997); Wittig, et al.,
20 Immunol. Letters (Netherlands), 57(1-3):217 (1997); and Oliveira and Odell, Oral Oncol.
(England), 33(4):260 (1997).

Extracellular and membrane-bound proteins play important roles in the formation,
differentiation and maintenance of multicellular organisms. The fate of many individual cells,
e.g., proliferation, migration, differentiation, or interaction with other cells, is typically
25 governed by information received from other cells and/or the immediate environment. This
information is often transmitted by secreted polypeptides (for instance, mitogenic factors,
survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which
are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins.
These secreted polypeptides or signaling molecules normally pass through the cellular

secretory pathway to reach their site of action in the extracellular environment, usually at a membrane-bound receptor protein.

Secreted proteins have various industrial applications, including use as pharmaceuticals, diagnostics, biosensors and bioreactors. In fact, most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane-bound proteins, also have potential as therapeutic or diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interaction. Membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. Transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly transmembrane proteins with homology to CD44 antigen. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.*, 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

SUMMARY OF THE INVENTION

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD44 antigen, wherein the polypeptide is designated in the present application as "PRO263".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO263 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO263 polypeptide having amino acid residues 1 to 322 of Figure 2 (SEQ ID NO:3), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides a vector comprising DNA encoding a PRO263 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host-cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO263 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO263 and recovering PRO263 from the cell culture.

5 In another embodiment, the invention provides isolated PRO263 polypeptide. In particular, the invention provides isolated native sequence PRO263 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 322 of Figure 2 (SEQ ID NO:3). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO263 polypeptide.

10 In another embodiment, the invention provides chimeric molecules comprising a PRO263 polypeptide or extracellular domain thereof fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO263 polypeptide fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to
15 a PRO263 polypeptide or extracellular domain thereof. Optionally, the antibody is a monoclonal antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

50064215-102997
20 Figures 1A and 1B show a nucleotide sequence (SEQ ID NO:1) containing the nucleotide sequence (SEQ ID NO:2) of a native sequence PRO263 cDNA, wherein the nucleotide sequence (SEQ ID NO:1) is a clone designated herein as "UNQ230" and/or "DNA34431-seq min". Also presented is the position of the initiator methionine residue, circled, at positions 1-3 of SEQ ID NO: 2 (positions 160-162 of SEQ ID NO: 1). The putative transmembrane domain of the protein is encoded by nucleotides beginning at nucleotide 709
25 of SEQ ID NO: 2 (underlined). The stop codon is also circled, immediately after the last nucleotide of SEQ ID NO: 2, at position 966.

Figure 2 shows the amino acid sequence (SEQ ID NO:3) derived from SEQ ID NO:

2.

30 Figures 3A and 3B show an alignment of nucleotide sequences from a variety of expressed sequence tags (SEQ ID NOS: 5-7) as well as a consensus nucleotide sequence derived therefrom designated "DNA30914-from dna" (SEQ ID NO: 4).

Figure 4 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the

CD44 antigen precursor from rat (SEQ ID NO: 8).

Figure 5 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of a metastasis-specific variant of the CD44 glycoprotein from rat (SEQ ID NO: 9).

Figure 6 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the CD44 protein from rat (SEQ ID NO: 10).

Figure 7 shows a BLAST sequence alignment analysis of portions of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with portions of the CD44 antigen precursor from mouse (SEQ ID NOS: 11 and 12).

Figures 8A and 8B show a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the CD44 antigen precursor from cricetus griseus (SEQ ID NO: 13).

Figure 9 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the CD44 antigen precursor from papio hamadryas (SEQ ID NO: 14).

Figure 10 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the CD44 antigen precursor from mesocricetus auratus (SEQ ID NO: 15).

Figure 11 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the human cell adhesion molecule CD44 (SEQ ID NO: 16).

Figure 12 shows a BLAST sequence alignment analysis of a portion of the PRO263 amino acid sequence derived from the DNA34431-seq min molecule with a portion of the human lymphocyte surface antigen precursor CDW44 (SEQ ID NO: 17).

Figure 13 shows two forward primers (SEQ ID NOS: 18 and 19), a reverse primer (SEQ ID NO: 20) and another primer (SEQ ID NO: 21).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO263 polypeptide" and "PRO263" when used herein encompass native sequence PRO263 and PRO263 polypeptide variants (which are further defined herein). The PRO263 polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

5 A "native sequence PRO263 polypeptide" comprises a polypeptide having the same amino acid sequence as a PRO263 polypeptide derived from nature. Such native sequence PRO263 polypeptide can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO263 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of a PRO263 polypeptide (e.g., an
10 extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of a PRO263 polypeptide. In one embodiment of the invention, the native sequence PRO263 polypeptide is a mature or full-length native sequence PRO263 polypeptide comprising amino acids 1 to 322 of Figure 2 (SEQ ID NO:3). In another embodiment of the invention, the native sequence PRO263 polypeptide is an
15 extracellular domain of the full-length PRO263 protein, wherein the putative transmembrane domain of the full-length PRO263 protein is encoded by nucleotides beginning at nucleotide 709 of SEQ ID NO: 2.

20 "Percent (%) amino acid sequence identity" with respect to the PRO263 amino acid sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO263 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using
25 publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

30 "Percent (%) nucleic acid sequence identity" with respect to the UNQ230 (DNA34387-seq min) sequence (SEQ ID NO:1) and the coding region therein, SEQ ID NO: 2, shown in Figures 1A and 1B is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the UNQ230 (DNA34387-seq min) sequence (SEQ ID NO:1) or SEQ ID NO:2, respectively, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining
percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal

alignment over the full length of the sequences being compared.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO263 polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO263 polypeptide-encoding nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO263 polypeptide-encoding nucleic acid. An isolated PRO263 polypeptide-encoding nucleic acid molecule is other than in the former setting in which it is found in nature. Isolated PRO263 polypeptide-encoding nucleic acid molecules therefore are distinguished from the PRO263 polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated PRO263 polypeptide-encoding nucleic acid molecule includes PRO263 polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express PRO263 polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably

linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO263 polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO263 antibody compositions with polypeptidic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Active" or "activity" for the purposes herein refers to form(s) of PRO263 which retain the biologic and/or immunologic activities of native or naturally-occurring PRO263 polypeptide. Variants of PRO263, have the activity of PRO263. Variants of particular interest are those wherein the molecule is truncated or there is a deletion therein so that there is no transdomain region.

II. Compositions and Methods of the Invention

A. Full-length PRO263 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO263. In particular, Applicants have identified and isolated cDNA encoding a PRO263 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO263 polypeptide have significant homology with the CD44 antigen and related proteins as shown in Figures 4-12. Accordingly, it is presently believed that PRO263 polypeptide disclosed in the present application is a newly identified member of the CD44 antigen family and possesses at least one of the properties associated with these antigens, *i.e.*, cancer and HIV marker, cell-cell or cell-matrix interactions, regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells.

In addition to the full-length native sequence PRO263 polypeptide described herein, it is contemplated that PRO263 variants can be prepared. PRO263 variants can be prepared by

introducing appropriate nucleotide changes into the PRO263-encoding DNA, or by synthesis of the desired PRO263 polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO263 polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

5 Variations in the native full-length sequence PRO263 or in various domains of the PRO263 polypeptide described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO263 polypeptide that results in a change in the amino acid sequence
10 of the PRO263 polypeptide as compared with the native sequence PRO263. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO263 polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO263 polypeptide with that of homologous
15 known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation
20 allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in the *in vitro* assay described in the Examples below.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed
25 mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO263-encoding variant DNA.

30 Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-

chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.]

5

B. Modifications of PRO263

Covalent modifications of PRO263 polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO263 polypeptide with an organic derivatizing agent that is capable of reacting with
10 selected side chains or the N- or C- terminal residues of a PRO263 polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO263 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO263 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with
15 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidyl)propionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and
20 lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO263 polypeptide included within the
25 scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO263 polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO263 polypeptide.

30 Addition of glycosylation sites to PRO263 polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO263 polypeptide (for O-linked glycosylation sites). The PRO263 amino acid sequence

may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO263 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO263 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO263 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO263 comprises linking the PRO263 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

PRO263 polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a PRO263 polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a PRO263 polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO263 polypeptide. The presence of such epitope-tagged forms of a PRO263 polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO263 polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a PRO263 polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu

HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

C. Preparation of PRO263

The description below relates primarily to production of PRO263 by culturing cells transformed or transfected with a vector containing at least nucleotides 1-966 of SEQ ID NO:2 of the UNQ230 (DNA34431-seq min) nucleic acid (SEQ ID NO:1). It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO263 polypeptides. For instance, the PRO263 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, L. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of PRO263 polypeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce a full-length PRO263 polypeptide.

1. Isolation of DNA Encoding PRO263

DNA encoding a PRO263 polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the PRO263 mRNA and to express it at a detectable level. Accordingly, human PRO263-encoding DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO263-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to a PRO263 polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may

be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO263 is to use PCR methodology [Sambrook et al., *supra*; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

- 5 The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., *supra*.

- 15 Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTAR, and INHERIT which employ various algorithms to measure homology.

- 20 Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., *supra*, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

- 25 Host cells are transfected or transformed with expression or cloning vectors described herein for PRO263 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., *supra*.

Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO₄,

and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO263-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism.

Suitable host cells for the expression of glycosylated PRO263 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells

(Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding the desired PRO263 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The desired PRO263 polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO263-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors

in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO263-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, **77**:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, **282**:39 (1979); Kingsman et al., Gene, **7**:141 (1979); Tschemper et al., Gene, **10**:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, **85**:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO263-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, **275**:615 (1978); Goeddel et al., Nature, **281**:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., **8**:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, **80**:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO263 polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., **255**:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., **7**:149 (1968); Holland, Biochemistry, **17**:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with

nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO263 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding a PRO263 polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO263 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO263.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO263 polypeptides in recombinant vertebrate cell culture are described in Gething et al., *Nature*, 293:620-625 (1981); Mantei et al., *Nature*, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription

of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO263 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO263-encoding DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO263 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO263 polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO263 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO263 polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO263 polypeptide produced.

D. Uses for PRO263

Nucleotide sequences (or their complement) encoding PRO263 polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO263-encoding nucleic acid will also be useful for the preparation of PRO263 polypeptides
5 by the recombinant techniques described herein.

The full-length nucleotide sequence SEQ ID NO:1 or the full-length native sequence PRO263 (SEQ ID NO:2) nucleotide sequence, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO263 gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of PRO263 or
10 PRO263 from other species) which have a desired sequence identity to the PRO263 sequence disclosed in Figures 1A and 1B (SEQ ID NOS: 1 or 2). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the UNQ230 (DNA34431-seq min) nucleotide sequence of SEQ ID NO:1 or SEQ ID NO: 2 as shown in Figures 1A and 1B or from genomic sequences including promoters, enhancer
15 elements and introns of native sequence PRO263-encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO263 gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ³²P or ³⁵S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling
20 systems. Labeled probes having a sequence complementary to that of the PRO263 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The probes may also be employed in PCR techniques to generate a pool of sequences
25 for identification of closely related PRO263 sequences.

Nucleotide sequences encoding a PRO263 polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO263 polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known
30 techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO263 encode a protein which binds to another protein (example, where the PRO263 polypeptide functions as a receptor), the PRO263

polypeptide can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO263 polypeptide can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO263 or a receptor for PRO263. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO263 polypeptide or any of its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO263 polypeptide can be used to clone genomic DNA encoding PRO263 in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO263. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO263 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO263 introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO263. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO263 can be used to construct a PRO263 "knock out" animal which has a defective or altered gene encoding PRO263 as a result of

homologous recombination between the endogenous gene encoding PRO263 and altered genomic DNA encoding PRO263 introduced into an embryonic cell of the animal. For example, cDNA encoding PRO263 can be used to clone genomic DNA encoding PRO263 in accordance with established techniques. A portion of the genomic DNA encoding PRO263 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO263 polypeptide.

PRO263 can be used in assays wherein CD44 antigen is generally used to determine PRO263 activity relative to that of CD44. The results can be used accordingly.

E. Anti-PRO263 Antibodies

The present invention further provides anti-PRO263 polypeptide antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO263 antibodies of the present invention may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections.

1
The immunizing agent may include the PRO263 polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete
5 adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

10 The anti-PRO263 antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to
15 the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO263 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell
20 line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably
25 contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

30 Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San

Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

5 The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against a PRO263 polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in
10 the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-
15 1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel
20 electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the
25 heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also
30 may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a

non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The anti-PRO263 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally,

5 a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

15 Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)).

20 4. Bispecific Antibodies

25 Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a PRO263 polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

30 Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen

combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving cross linking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

F. Uses for anti-PRO263 Antibodies

The anti-PRO263 antibodies of the present invention have various utilities. For example, anti-PRO263 antibodies may be used in diagnostic assays for PRO263 polypeptides, e.g., detecting expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the

antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO263 antibodies also are useful for the affinity purification of PRO263 polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a PRO263 polypeptide are immobilized on a suitable support, such as a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO263 polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO263 polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO263 polypeptide from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1: Isolation of cDNA Clones Encoding Human PRO263

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with

the program "phrap" (Phil Green, University of Washington, Seattle, Washington; <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA30914-from dna (see Figures 3A and 3B).

5 Based on the DNA30914-from dna consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO263. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest by the *in vivo* cloning procedure using the probe oligonucleotide and one of the primer pairs.

15 PCR primers were synthesized to form two forward primers (SEQ ID NOS: 18 and 19) and one reverse primer (SEQ ID NO: 20) as shown in Figure 13. Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30914-from dna sequence which has the sequence of SEQ ID NO: 21 also shown in Figure 13.

20 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO263 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

25 The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, **253**:1278-1280 (1991)) in the unique XhoI and NotI sites.

30 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO263 [herein designated as UNQ230 (DNA34431-seq min)] (SEQ ID NO:1)

and the derived protein sequence for PRO263.

The entire nucleotide sequence of UNQ230 (DNA34431-seq min) is shown in Figures 1A and 1B (SEQ ID NO:1). Clone UNQ230 (DNA34431-seq min) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 1-3 [Kozak et al., *supra*] of SEQ ID NO: 2 and ending at the stop codon after the nucleotide at position 5 966 of SEQ ID NO: 2 (Figures 1A and 1B). The predicted polypeptide precursor is 322 amino acids long (Figure 2). Clone UNQ230 (DNA34431-seq min) has been deposited with ATCC and is assigned ATCC deposit no. 209399.

Analysis of the amino acid sequence of the full-length PRO263 polypeptide suggests that portions of it possess significant homology to CD44 antigen as shown in Figures 4-12, 10 thereby indicating that PRO263 may be a novel cell surface adhesion molecule.

EXAMPLE 2: Use of PRO263-encoding DNA as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO263 as a hybridization probe.

15 DNA comprising the coding sequence of full-length PRO263 (as shown in Figures 1A and 1B, SEQ ID NO: 2 or 1) is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO263) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under 20 the following high stringency conditions. Hybridization of radiolabeled UNQ230 (DNA34431-seq min)-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

25 DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO263 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO263 Polypeptides in *E. coli*

30 This example illustrates the preparation of unglycosylated forms of PRO263 polypeptides by recombinant expression in *E. coli*.

The DNA sequence encoding PRO263 (SEQ ID NO: 3) is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond

to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO263 coding region, lambda transcriptional terminator, and an argU-gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO263 polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the polypeptide.

EXAMPLE 4: Expression of PRO263 Polypeptides in Mammalian Cells

This example illustrates preparation of glycosylated forms of PRO263 polypeptides by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO263-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO263-encoding DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO263.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO263 DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES

(pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

- 5 Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ³⁵S-cysteine and 200 μ Ci/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of
- 10 PRO263 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO263-encoding DNA may be introduced into 293 cells transiently using the dextran sulfate method described by Sompariyac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g

15 pRK5-PRO263 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the

20 conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO263 polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

- In another embodiment, PRO263 polypeptide can be expressed in CHO cells. The pRK5-PRO263 vector can be transfected into CHO cells using known reagents such as Ca PO₄
- 25 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO263 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO263
- 30 polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO263 polypeptide may also be expressed in host CHO cells. The PRO263-encoding DNA may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a

Baculovirus expression vector. The poly-his tagged PRO263-encoding DNA insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO263 polypeptide can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

EXAMPLE 5: Expression of a PRO263 Polypeptide in Yeast

The following method describes recombinant expression of PRO263 polypeptides in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO263 polypeptide from the ADH2/GAPDH promoter. DNA encoding the PRO263 polypeptide of interest, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO263 polypeptide. For secretion, DNA encoding the PRO263 polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO263 polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO263 polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO263 polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 6: Expression of PRO263 Polypeptides in Baculovirus

The following method describes recombinant expression of PRO263 polypeptides in Baculovirus.

The PRO263-encoding DNA is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids

derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO263-encoding DNA or the desired portion of the PRO263-encoding DNA (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharming) into *Spodoptera frugiperda* ("Sf9") cells (ATCC-CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 to 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications.

Viral infection and protein expression is performed as described by O'Reilley et al., *Baculovirus expression vectors: A laboratory Manual*, Oxford:Oxford University Press (1994).

Expressed poly-his tagged PRO263 polypeptide can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO263 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO263 polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 7: Preparation of Antibodies that Bind PRO263 Polypeptides

This example illustrates the preparation of monoclonal antibodies which can specifically bind to PRO263 polypeptides.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO263 polypeptide, fusion proteins containing a PRO263 polypeptide, and cells expressing recombinant PRO263 polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO263 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO263 polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO263 polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO263 polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against a PRO263 polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO263 polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively,

affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Deposit of Material

The following materials have been deposited with the American Type Culture
5 Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
Plasmid	209399	October 17, 1997

10 This deposit was made under the provisions of the Budapest Treaty on the International
Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the
Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the
deposit for 30 years from the date of deposit. The deposit will be made available by ATCC
under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc.
and ATCC, which assures permanent and unrestricted availability of the progeny of the culture
15 of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to
the public of any U.S. or foreign patent application, whichever comes first, and assures
availability of the progeny to one determined by the U.S. Commissioner of Patents and
Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules
pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

20 The assignee of the present application has agreed that if a culture of the materials on
deposit should die or be lost or destroyed when cultivated under suitable conditions, the
materials will be promptly replaced on notification with another of the same. Availability of
the deposited material is not to be construed as a license to practice the invention in
contravention of the rights granted under the authority of any government in accordance with
25 its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled
in the art to practice the invention. The present invention is not to be limited in scope by the
construct deposited, since the deposited embodiment is intended as a single illustration of
certain aspects of the invention and any constructs that are functionally equivalent are within
30 the scope of this invention. The deposit of material herein does not constitute an admission
that the written description herein contained is inadequate to enable the practice of any aspect
of the invention, including the best mode thereof, nor is it to be construed as limiting the scope
of the claims to the specific illustrations that it represents. Indeed, various modifications of

the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

60064215-102997

WHAT IS CLAIMED IS:

1. Isolated nucleic acid comprising DNA encoding a PRO263 polypeptide having amino acid residues 1 to 322 of Figure 2 (SEQ ID NO:3).

2. The nucleic acid of Claim 1 wherein said DNA comprises the nucleotide sequence of SEQ ID NO: 2 or its complement.

3. The nucleic acid of Claim 1 wherein said DNA comprises the nucleotide sequence of SEQ ID NO: 2 or its complement.

4. An isolated nucleic acid comprising the nucleotide sequence of the full-length coding sequence of clone UNQ230 (DNA34431) deposited under accession number ATCC 209399.

5. An isolated nucleic acid encoding an extracellular domain of a PRO263 polypeptide.

6. A vector comprising the nucleic acid of any one of Claim 1, Claim 4 or Claim 5.

7. The vector of Claim 6 operably linked to control sequences recognized by a host cell transformed with the vector.

8. A host cell comprising the vector of Claim 6.

9. The host cell of Claim 8 wherein said cell is a CHO cell.

10. The host cell of Claim 8 wherein said cell is an *E. coli*.

11. The host cell of Claim 8 wherein said cell is a yeast cell.

12. A process for producing a PRO263 polypeptide comprising culturing the host cell of Claim 8 under conditions suitable for expression of said PRO263 polypeptide and recovering said PRO263 polypeptide from the cell-culture.

13. Isolated native sequence PRO263 polypeptide comprising amino acid residues 1 to 322 of Figure 2 (SEQ ID NO:3).

14. An isolated extracellular domain of a PRO263 polypeptide.

5 15. A chimeric molecule comprising a PRO263 polypeptide fused to a heterologous amino acid sequence.

16. The chimeric molecule of Claim 15 wherein said heterologous amino acid sequence is an epitope tag sequence.

10

17. The chimeric molecule of Claim 15 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

18. An antibody which specifically binds to a PRO263 polypeptide.

15

19. The antibody of Claim 18 wherein said antibody is a monoclonal antibody.

60064215-102297

Abstract of the Disclosure

The present invention is directed to novel polypeptides having homology to CD44 antigen and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods
5 for producing the polypeptides of the present invention.

60064215-102997

PO1177

```
>< /usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA34431 (2554 bases)
< good sequence: 3-2554 (2552 bases)
< insert: 84-2455 (2372 bases), 9 regions found
< 5' PRK5 + lk: 17-901, 65 matches (97%), 28 consec, 2 gaps
< 3' PRK5 + lk: 2456-1566, 66 matches (100%), 66 consec
< PRK5D: 2458-1366, 82 matches (97%), 76 consec, 2 gaps
< PRK5D: 1-885, 37 matches (90%), 30 consec, 2 gaps
< PRK5D: 22-914, 30 matches (85%), 16 consec, 1 gap
< PRK5B: 2458-1357, 82 matches (97%), 76 consec, 2 gaps
< PRK5B: 1-885, 69 matches (97%), 43 consec, 2 gaps
< 3' cDNA linker: 2456-1566, 26 matches (100%), 26 consec
< 5' cDNA linker: 68-950, 16 matches (100%), 16 consec
>< [DNA34431],
AGCAGGGAATCCGGATGTCCTGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATA
GTTCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCCTCTCATATACACAGTGGC
CATCTGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG
<MET (trans=1-s,dir=f,res=1)
ATGSCCAGGTGCTTCAGCCTGGTGTGCTTCTCATTCCATCTGGACCACGAGGCTCCTTG
GTTCCAGGCTCTTTGCGTGACAGAGAGCTTCCATCCAGGTGTCAATGAGAATATGGGG
ATCACCCCTGTGAGCAAAAGGCGAACACAGCAGTGAATTTACAGAACTAAGGAGGCC
TGATGGCTGCTGGGACTAAAGTTTGGCCGGCAAGGACCAAGTTGAACAGCCTTGAAAGCT
AGCTTTGAAACTTTGCAGCTATGGCTGGTGGAGATGGATCGTGGTCACTCTAGGATT
AGCCCAAACCCCAAGTGTGGGAAAATGGGGTGGGTGCTCTGATTTGGAGGTTCCAGTG
AGCCGACAGTTTGAGCCTTATGTTACAACTCATCTGATCTTGGACTAATCTCGTGCAAT
CCAGAAATATACACCAAGATCCCATATTTCAACACTCAAACTGCAACACAAACAACA
GAATTTATTTGTCAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCC
CCTACTACTACTCTCTCTGCTCCAGCTTCCACTTCTATTTCCAGGAGAAAATAATTGATT
TGTGTACACAGAGTTTATGGAACCTAGCACCATGTCTACAGAACTGAACCATTTGTT
GAAATAAAGCAGCATTTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCACCGGCTCTG
CTAGTGCTTGTCTCTCTCTTTGTTGGTGTGTCAGCTGGTCTTGGATTGTGTATGTCAAA
AGGTATGTGAAGGCCTTCCCTTTTACAAACAGAAATCAGCAGAGGAAATGATCGAAACC
AAAGTAGTAAAGGAGGAGAGGCCCAATGATAGCAACCTTAATGAGGAATCAAGAAACT
```

3000 NO: 1

10 NO: 2

FIGURE 1A

60064215.102997

GATAAAAACCCAGAAGAGTCCAGAGTCCAAGAGTCCAAGCAAAACTACCGTGGGATGCCCTGGAAGCT
GAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGGTTTCCTTCATGCTCCT
TACCCCTGCCCCAGCTGGGAAATCAAAAGGGCCAAAGAACCAAGAAAGTCCACCCCT
TGGTTCCTAACTGGAATCAGCTCAGGACTGCCATTTGGACTATGGAGTGCACCAAGAGAA
TGCCCTTTCCTTATTTGTAACCTTGTCTGGATCCTATCCTCCTACCCTCCAAAGCTTTCCCA
CGCCCTTTCCTAGCCTGGCTATGTCCTTAATAATATCCACTGGGAGAAAGGAGTTTTCGCA
AGTGCAAGGACCTAAACAATCTCATCAGTATCCAGTGGTAAAAAGGCCCTCTACTGATTC
GAGGCTAGGTGGGTGAAAGCCCAAGGAGTCACTGAGACCAAGGCTTTCTACTGATTC
GCAGCTCAGACCCCTTTCCTCAGCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCT
TCTGAGCCCGGTAGAGCAAAAGAAATGGCAGAAAGTTTAGCCCCCTGAAAGCCATGGAGA
TCTCATAACTTGAGACCTAATCTCTGTAAAGCTAAATAAGAAATAGAAACAAGGCTGA
GGATACGACAGTACACTGTGAGCAGGACGTGTAACACACAGACAGGCTCAAAGTGTTCCT
CTGAACACATTTGAGTTGGAAATCAGTGTGTTTAGAACACACACACTTACTTTTCTGGTCTCT
ACCACTGCTGATATTTCTCTAGGAAATATATCTTTTACAGTAACAAAATAAAACTCT
TATAAATTTCTATTTTATCTGAGTTACAGAAATGATTTATTAAGGAAGATTAATCAGTAA
TTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTCAAAGT
GCTGTCAAGGTATNACACTCTGTAAATGAATATTTATTTCTCAAAAATTGCACATAGTA
GAACGCTATCTGGGAAGCTATTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAG
ACTAATTTTATTTTTCGTGAGACTAATCTTAATTCATTTTCTCTAAATATGGCAACCATTA
TAACCTTAATTTTATTAATAACATACCTAAGAAATGATGTTTACNTATATATACCAAGC
ACATTTTAAAGTGCCATTAAACAAATGATCAGTACCCCTCTTTTTCACCAAGAGGG
ACTGAGAGATGCAGAAATATTTGTGACAAAATAATTAAGCATTTAGAAAACTT

F1602E 1B

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34431
><subunit 1 of 1, 322 aa, 0 stop
><MW: 35213, pI: 8.71, NX(S/T): 3
MARCFSLVLLLTSTWTRLLVQGSILRAEELSIQVSCRIMGITLVSKKANQQLNFTAEKEA
CRLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGLIWKVPV
SRQFAAYCYNSSDIWTNSCIPETITTKDPIFNTQTATQTFEFIVSDSTYSVASPYSTIPA
PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVP
LVLALLFFGAAGLGFYVYVYKAFPFITKNQKEMIETKVVKKEKANDSNPNEESKKT
DKNPEESKSPSKTTVRCLEAEV **SEQ ID No: 3**

FIGURE 2

<DNA30914: 3 members (1 incyte, 2 est) 623 bp, 4 gaps, 306 bp orf (+3)

13040: S AA081350
ADD0: C 1417608
3D00: 4 <DNA30914>
1 GGCTCCTGGTCCCAAGGCTCTTTGCGTGCAGAAAGAGCTTTCCATCCAGGTG G
1
1 GGCTCCTGGTCCCAAGGCTCTTTGCGTGCAGAAAGAGCTTTCCATCCAGGTG
51 TCATGCAGAAATTATGGGGATCACCCCTTGTGAGCAAAAAGGGCAACCAGCA
2 TCATGCAGAAATTATGGGG-TCACC-TTGTGAGCAAAAAGGGCAACCAGCA
1 GGATCACCCCTTGTGAGCAAAAAGGGCAACCAGCA
51 TCATGCAGAAATTATGGGGATCACCCCTTGTGAGCAAAAAGGGCAACCAGCA
101 GCTGAATTTTCACAGAAGCTAAGGAGGCCCTGTAGN-TGCTGGGACTAAGTT
52 GCTGAATTTTCACAGAAGCTAAGGAGGCCCTGTAGGCTGCTGGGACTAAGTT
35 GCTGAATTTTCACAGAAGCTAAGGAGGCCCTGTAGG-TGCTGGGACTAAGTT
101 GCTGAATTTTCACAGAAGCTAAGGAGGCCCTGTAGGCTGCTGGGACTAAGTT
151 TGGCCGGCAAGGACCAAGTTGAAACAGCCCTTGAAAGCTAGCTTTGAAACT
102 TGGCCGGCAAGGACCAAGTTGAAACAGCCCTTGAAAGCTAGCTTTGAAACT
85 TGGCCGGCAAGGACCAAGTTGAAACAGCCCTTGAAAGCTAGCTTTGAAACT
151 TGGCCGGCAAGGACCAAGTTGAAACAGCCCTTGAAAGCTAGCTTTGAAACT
201 TGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTCCTAGGATTA
152 TGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTC-TAGGAT-A
135 TGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTC-TAGGATTA
201 TGCAGCTATGGCTGGGTTGGAGATGGATTTCGTGGTCATCTC-TAGGATTA
251 GCCCAAAACCCCAAAGTGTGGGAAAAAATGGGGTGGGTCTCTGATTTTGG
202 GCCCAAAACCCCAA-GTGTGGGAAAAA-TGGGGT
185 GCCCAAAACCCCAA-GTGTGGGAAAAA-TGGGGTGGGTCTCTGATTTTGG
251 GCCCAAAACCCCAA-GTGTGGGAAAAA-TGGGGTGGGTCTCTGATTTTGG
301 AAGGTTCCAAAGTGAACCGACAAAGTTTTCAGCCCTAATTGTTACAACATCAT
235 AAGGTTCCA-GTGAGCCGACA-GTTT-GCAGCCTA-TTGTACAACATCAT
301 AAGGTTCCAAGTGAACCGACAAAGTTTTCAGCCCTAATTGTTACAACATCAT
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>
AA081350
1417608
AA046671
<DNA30914>

FIGURE 3A

150054215.102997

AA081350	351	CTGATACNTGGGACTAACTCGTGCCATTCCA
AA046671	285	CTGATACTTGG-ACTAACTCGTGCATTCCAGAAATTATCACCACCAAGA
<DNA30914>	351	CTGATACTTGG-ACTAACTCGTGCATTCCAGAAATTATCACCACCAAGA
AA046671	335	TCCCATATTCAACACTCAAACTGCAACACAAACAAGAAATTTATTGTCA
<DNA30914>	401	TCCCATATTCAACACTCAAACTGCAACACAAACAAGAAATTTATTGTCA
AA046671	385	GTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCT
<DNA30914>	451	GTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCT
AA046671	435	ACTACTACTCCTCCTGCTCCAGTTCACCTTCTATTCCACGGAGAAAAAAA
<DNA30914>	501	ACTACTACTCCTCCTGCTCCAGTTCACCTTCTATTCCACGGAGAAAAAAA
AA046671	485	TTGATTGTGTACAGAAAGTTTTTATGGAAACTAGCACCATGTCTACAGA
<DNA30914>	551	TTGATTGTGTACAGAAAGTTTTTATGGAAACTAGCACCATGTCTACAGA
AA046671	535	AACTGAACCATTTGTTGAAAAATAAGCA
<DNA30914>	601	AACTGAACCATTTGTTGAAAAATAAGCA

Figure 3B

1	CD44_RAT	Cd44 antigen precursor - rattus norvegicus	+1	217	52	32
2	P_R14768	Metastasis-specific variant of CD44 glycoprotein	+1	217	52	32
3	RNU96138_1	CD44 protein - Rattus norvegicus	+1	217	52	32
4	CD44_MOUSE	Cd44 antigen precursor - mus musculus	+1	208	52	33
5	CD44_CRIGR	Cd44 antigen precursor - cricetus griseus	+1	204	57	34
6	CD44_PAPHA	Cd44 antigen precursor - papio hamadryas	+1	202	59	33
7	P_R07355	B7 adhesion receptor - Papio sp.	+1	202	59	33
8	CD44_MESAU	Cd44 antigen precursor - mesocricetus aur...	+1	202	56	33
9	JH0417	cell adhesion molecule CD44 - human	+1	200	54	32
10	A32376	lymphocyte surface antigen precursor CDW44 -	+1	200	54	32

>1 CD44_RAT Cd44 antigen precursor - rattus norvegicus (503 aa)
Score = 217 (76.4 bits), Expect = 1.2e-13, P = 1.2e-13
Identities = 52/163 (31%), Positives = 77/163 (47%), at 202,10, Frame = +1

```

DNA34431 202 WTTRLLVQGSRLRAEELSIQVSCRIMGITLVSKKANQQLNFTAEKACRLLGLSLAGKQOV
          * * * * * . . . . . * * * * * * * * * * * * * * * *
CD44_RAT 10 WGLLCILLQLSLAQQQIDILNITCRYAGVFHVEKNGRYSISRTEAADLCEAFNTTLPMTAQM
          * * * * * . . . . . * * * * * * * * * * * * * * * *
DNA34431 382 ETALKASFETCSYGVWGDGFVWISRISPNPKCGKNGVGLIWKVPVSRQFAAYCYNSSDT
          * * * * * . . . . . * * * * * * * * * * * * * * * *
CD44_RAT 70 ELALRKGFETCRYGFI-EGHVVIPRIHPNAICAANNITGVYILLASNTSHYDITYCFNASAP
          * * * * * . . . . . * * * * * * * * * * * * * * * *
DNA34431 562 WTNSCIPEIITTKDPIFNQTATQTTEFIVSDST-YSVASPYST
          * * * * * . . . . . * * * * * * * * * * * * * * * *
CD44_RAT 129 LEEDC-----TSVTDLENSFDGPFVTITIVNRDGTTRYSKKGEYRT
          * * * * * . . . . . * * * * * * * * * * * * * * * *

```

FIGURE 4

>2 P_R14768 Metastasis-specific variant of CD44 glycoprotein - Rattus (503 aa)
 Score = 217 (76.4 bits), Expect = 1.2e-13, P = 1.2e-13
 Identities = 52/163 (31%), Positives = 77/163 (47%), at 202,10, Frame = +1

```

DNA34431 202 WTTRLLVQGSRLRAEELSIQVSCRIMGITLVSKKANQQLNFTAEACRLLGLSLAGKDQV
      * * * * * . . . . . * * * * * . . . . . * * * * *
P_R14768 10 WGLLCLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEADLCEAFNTTLPPTMAQM
      * * * * * . . . . . * * * * * . . . . . * * * * *
DNA34431 382 ETALKASFETCSYGVWGDGFVVISRISPNPKCGKNGVGLTWKVPVSRQFAAYCYNSSDT
      * * * * * . . . . . * * * * * . . . . . * * * * *
P_R14768 70 ELALRKGFETCRYGFI-EGHVIPRIHPNAICAANNITGVYILLASNTSHYDTYCFNASAP
      * * * * * . . . . . * * * * * . . . . . * * * * *
DNA34431 562 WTNSCIPEIITTKDPIFNQTQTATQTFEFIVSDST-YSVASPYST
      * * * * * . . . . . * * * * * . . . . . * * * * *
P_R14768 129 LEEDC-----TSVTDLPSNFDGPPVTITIVNRDGTTRYSKKGEYRT

```

FIGURE 5

>3 RNU96138_1 CD44 protein - Rattus norvegicus (780 aa)
 Score = 217 (76.4 bits), Expect = 2.9e-13, P = 2.9e-13
 Identities = 52/163 (31%), Positives = 77/163 (47%), at 202,10, Frame = +1

```

DNA34431 202 WTTRLLVQGSRLRAEELSIQVSCRIMGITLVSKKANQQLNFTAEACRLLGLSLAGKDQV
      * * * * * . . . . . * * * * * . . . . . * * * * *
RNU96138_1 10 WGLLCLQLSLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEADLCEAFNTTLPPTMAQM
      * * * * * . . . . . * * * * * . . . . . * * * * *
DNA34431 382 ETALKASFETCSYGVWGDGFVVISRISPNPKCGKNGVGLTWKVPVSRQFAAYCYNSSDT
      * * * * * . . . . . * * * * * . . . . . * * * * *
RNU96138_1 70 ELALRKGFETCRYGFI-EGHVIPRIHPNAICAANNITGVYILLASNTSHYDTYCFNASAP
      * * * * * . . . . . * * * * * . . . . . * * * * *
DNA34431 562 WTNSCIPEIITTKDPIFNQTQTATQTFEFIVSDST-YSVASPYST
      * * * * * . . . . . * * * * * . . . . . * * * * *
RNU96138_1 129 LEEDC-----TSVTDLPSNFDGPPVTITIVNRDGTTRYSKKGEYRT

```

FIGURE 6

```
>4 CD44_MOUSE      Cd44 antigen precursor - mus musculus (655 aa)
Score = 208 (73.2 bits), Expect = 2.2e-12, Sum P(2) = 2.2e-12
Identities = 52/158 (32%), Positives = 77/158 (48%), at 217,14, Frame = +1

DNA34431          217 LVQGSRLRAEELSISIQVSCRIMGITLVSKKANQQQLNFTEAKEACRLLGLSLAGKDOVETALK
                  *.* ** .....*** ** * ..*** .*. * *** **
ID#2: 11 CD44_MOUSE      14 LQLSLAHQQIDLNVTCRYAGVFHVEKNGRYSISIRTEAADLCOAFNSTLPITMDQMKLALS
                  *.* ** .....*** ** * ..*** .*. * *** **
DNA34431          397 ASFETCSYGWVGDFVVISRI SPNPKCKNGVGVLWKVPVSQRQFAAYCYNSSDTWTWNSC
                  **** *..* *** ** * * * * * * * * * * * * * * * *
CD44_MOUSE        74 KGFETCRYGFI-EGNVVIPRIHPNAICAAHTGVYILVTSTSHYDTCYCFNASAPPEEDC
                  **** *..* *** ** * * * * * * * * * * * * * * * *

DNA34431          577 IPELIITTKDPIFNITQTATQTTFEFIVSDST-YSVASPYST
                  *.....*..* * * * * * * * * * * * * * * *
CD44_MOUSE        133 -----TSVTDLPNSFDGPVTITIVNRDGGTRYSKKGTYRT
                  *.....*..* * * * * * * * * * * * * * * *

Score = 48 (16.9 bits), Expect = 2.2e-12, Sum P(2) = 2.2e-12
Identities = 12/39 (30%), Positives = 22/39 (56%), at 1006,604, Frame = +1

DNA34431          1006 VKEEKANDSNPNEESKKTTDKNPESKSPSKTTVRCLEAE
                  *..* * * * * * * * * * * * * * * *
ID#2: 12 CD44_MOUSE     604 VEDRKRPSELN-GEASKSQEMVHLVNKEPSETPDQCMTAD
                  *..* * * * * * * * * * * * * * * *
```

5 CD44 antigen precursor - cricetus griseus (362 aa)
 Score = 204 (71.8 bits), Expect = 1.2e-12, p = 1.2e-12
 Identities = 57/169 (33%), Positives = 80/169 (47%), at 229, 18, Frame = +1

DNA34431	229	SLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLGLSLAGKQDQVETALKASFE
		** * ** * * * * ** * * * * ** * * *
CD44_CRIGR	18	SLAHEQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMDQMVALSKGFE
DNA34431	409	TCSYGVWVGDFVWISRI SPNPKCGKNGVGLIWKVPVSRQFAAYCYNSSDTWNSCIPEI
		** * * * . . *
CD44_CRIGR	78	TCRYGFI-EGHWIPRIQPNICAANHGTGVYILTNTSH-YDTYCFNASAPLEEDC-----

>6 CD44_PAPHA CD44 antigen precursor - papio hamadryas (362 aa)
 Score = 202 (71.1 bits), Expect = 2.1e-12, P = 2.1e-12
 Identities = 59/180 (32%), Positives = 84/180 (46%), at 217,14, Frame = +1

```

DNA34431 217 LVQGSRLRAEELSIOVSCRIMGITLVSKKANQQLNFTTEAKEACRLLGLSLAGKDQVETALK
*** ** .. ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **
EQ DND:WCD44_PAPHA 14 LVQLSLA--QIDLNITCRFEGIYHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALS
          . . . . .
DNA34431 397 ASFETCSYGVWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDITWNSC
**** ** . * *** ** ** * * * * * * * * * * * * * * * *
CD44_PAPHA 72 IGFETCRYGFI-EGHVVIPRIHPNSICAANNITGVYILTSNTS-QYDTCFNASAPPGEDC
          . . . . .
DNA34431 577 --IPEIITTKD-PIFNITQTATQTTEFIVSDSYSVASPYSTIPAPTITPPAPASTSIPRR
          * ** * * * * * * * * * * * * * * * * * * * *
CD44_PAPHA 130 TSVTDLPNAFDGPITITIVNRDGTTRY-VKKGEYRT-NPEDINPSSPTDDDDVSSGSSSERS
          . . . . .
DNA34431 748 KKL
          *
CD44_PAPHA 188 STL
  
```

FIGURE 9

8 CD44_MESAU Cd44 antigen precursor - mesocricetus auratus (431 aa)
 Score = 202 (71.1 bits), Expect = 3.7e-12, P = 3.7e-12
 Identities = 56/169 (33%), Positives = 80/169 (47%), at 229, 18, Frame = +1

DNA34431 229 SLRAEELSIQVSCRIMGITLVSKKANQQLNFTTEAKEACRLGLSLAGKQQVETALKASFEE
 ** ** * * * * * ** * * * * * ** * * * * * ** * * * * * ** * * * * *
 CD44_MESAU 18 SLAQQQIDLNITCRYAGVFHVEKNGRYSISRTEAADLCQAFNSTLPTMDQMWMAALSKGFE

DNA34431 409 TCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQFAAYCYNSSDWTNSCIPEI
 **
 CD44_MESAU 78 TCRYGFI-EGHVVIPRIQPNICAANHHTGVYILTSNTSH-YDTYCFNASAPLEEDC-----

DNA34431 589 ITTKDPPIENTQTATQTTEFIVSDST-YSVASPYST-----IPAPTTTTPPAPASTS
 *
 CD44_MESAU 132 -TSVTDLPNSEFEGPVTTITIVNRDGTTRYKKGEYRTHQEDIDASNTTDDDDVSSGS

FIGURE 10

>9 JH0417 cell adhesion molecule CD44 - human (361 aa)
 Score = 200 (70.4 bits), Expect = 3.5e-12, P = 3.5e-12
 Identities = 54/168 (32%), Positives = 80/168 (47%), at 232,17, Frame = +1

```

DNA34431 232 LRAEELSIQVSCRIMGITLVSKKANQQLNFTFAKEACRLGLSLAGKQDQVETALKASFET
      * .. . . . * * * * * * * * * * * * * * * * * * * * * * *
JH0417 17 LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET
      * .. . . . * * * * * * * * * * * * * * * * * * * * * * *

DNA34431 412 CSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYNSSDIWTNSC--IPE
      * * * * * * * * * * * * * * * * * * * * * * * * * * * *
JH0417 77 CRYGFI-EGHVVIPIRIHPNSICAANNITGVYILTYNTS-QYDTYCFNASAPPEEDCTSVTD
      * * * * * * * * * * * * * * * * * * * * * * * * * * * *

DNA34431 586 IITTKD-PIFNTQTATQTTEFIVSDSTYSVASPYSTIPA-PTTTPPAPASTS
      * * * * * * * * * * * * * * * * * * * * * * * * * * * *
JH0417 135 LPNAFDGPITITIVNRDQTRY-VQKGEYRT-NPEDIYPSNPTDDDDVSSGSSS
      * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

SeqID 40:16

FIGURE 11

>10 A32376 lymphocyte surface antigen precursor CDW44 - human (361 aa)
 Score = 200 (70.4 bits), Expect = 3.5e-12, P = 3.5e-12
 Identities = 54/168 (32%), Positives = 80/168 (47%), at 232,17, Frame = +1

DNA34431	232	LRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLGLSLAGKDQVETALKASFET
A32376	17	LSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFET
DNA34431	412	CSYGMVGDGFVVISRISPNPKCGKNGVGLIWKVPVSRQFAAYCYNSSDTWTNSC--IPE
A32376	77	CRYGFI-EGHVIPRIHPNSICAANNITGVVILTYNTS-QYDTYCFNASAPPEEDCTSVTD
DNA34431	586	IITTKD-PIFNTQTATQTTFEIVSDSTYSVASPYSTIPA-PTTTPPAPASTS
A32376	135	LFNAFDGPITITIVNRDGTTRY-VQKGEYRT-NPEDIYPSNPTDDDDVSSGSSS

FIGURE 12

30914.P

AGGAGGCCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTT SEQ ID NO: 21

.f1

GAGCTTTCCATCCAGGTGTCAATGC SEQ ID NO: 18

.r

TGGAGCAGGAGGAGTAGTAGTAGG SEQ ID NO: 20

.f2

GTCAGTGACAGTACCTACTCGG SEQ ID NO: 19

FIGURE 13

60064215-102997



PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

EXPRESS MAIL MAILING LABEL

Number: 224 104 541 063 US Date of Deposit: April 22, 1998

I hereby certify that this paper or file is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to Box Provisional Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231, on April 22, 1998

Typed or Printed Name: Kevin Campbell

Signed: *K. Campbell*

Docket Number:	P-65410/WH/MTK Genentech Docket No. PR1189	Type a plus sign (+) inside this box -	+
----------------	---	--	---

INVENTOR(S)/APPLICANT(S)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Wood	William		San Mateo, California
Goddard	Audrey		San Francisco, California
Gurney	Austin		Belmont, California
TITLE OF THE INVENTION (280 characters max)			
NOVEL POLYPEPTIDES HAVING HOMOLOGY TO DIFF33 AND NUCLEIC ACIDS ENCODING THEREFOR			
CORRESPONDENCE ADDRESS			
Walter H. Dreger, Registration No. 24, 190 FLEHR HOBBACH TEST ALBRITTON & HERBERT LLP, 4 Embarcadero Center, Suite 3400, San Francisco			
STATE	CA	ZIP CODE	COUNTRY
		94111	US
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/>	Specification	Number of Pages	38
<input checked="" type="checkbox"/>	Drawings	Number of Sheets	38
<input type="checkbox"/>			Small Entity Statement
<input type="checkbox"/>			Other (specify):
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)			
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the Provisional filing fees		PROVISIONAL FILING FEE AMOUNT (\$)
<input type="checkbox"/>	The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number:		
	06-1300 (Order No. P-65410/WH/MTK)		\$150/\$75
			\$150

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒

No.

☐

Yes, the name of the U.S. Government Agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE: *Mark T. Kresnak*

Date

April 22, 1998

TYPED or PRINTED NAME Mark T. Kresnak, Ph.D.

REGISTRATION NO.
(if appropriate)

P-42,767

☐

Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take 2 hours to complete. Times will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Assistant Commissioner for Patents, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

585617



65705 U.S. PTO
60/082797
04/22/98

PATENT
Attorney Docket No.: P-65410/WHD/MTK
(Genentech Docket No.: PR1189)

NOVEL POLYPEPTIDES HAVING HOMOLOGY TO DIFF33
AND NUCLEIC ACIDS ENCODING THEREFOR

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides having homology to the human Diff 33 protein, designated herein as "PRO300" polypeptides.

BACKGROUND OF THE INVENTION

Membrane-bound proteins and receptors can play an important role in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interaction. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

Efforts are being undertaken by both industry and academia to identify new, native receptor proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins.

The Diff 33 protein is over-expressed in mouse testicular tumors. At present its role is unclear, however, it may play a role in cancer. Given the medical importance of understanding the physiology of cancer, efforts are currently being under taken to identify new, native proteins which are involved in cancer. We describe herein the identification of a novel polypeptide which has homology to Diff 33.

SUMMARY OF THE INVENTION

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO300".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO300 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO300 polypeptide having amino acid residues 1 to 457 of Figure 2 (SEQ ID NO:3), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. The isolated nucleic acid sequence may comprise the cDNA insert of the DNA40625-1189 vector deposited on April 21, 1998 as ATCC _____ which includes the nucleotide sequence encoding PRO300.

In another embodiment, the invention provides a vector comprising DNA encoding a PRO300 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO300 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO300 and recovering PRO300 from the cell culture.

In another embodiment, the invention provides isolated PRO300 polypeptide. In particular, the invention provides isolated native sequence PRO300 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 457 of Figure 2 (SEQ ID NO:3). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO300 polypeptide. Optionally, the PRO300 polypeptide is obtained or is obtainable by expressing the polypeptide encoded by the cDNA insert of the DNA40625-1189 vector deposited on April 21, 1998 as ATCC _____.

In another embodiment, the invention provides chimeric molecules comprising a PRO300 polypeptide or extracellular domain thereof fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO300 polypeptide fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to a PRO300 polypeptide or extracellular domain thereof. Optionally, the antibody is a monoclonal antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) containing the nucleotide sequence (SEQ ID NO:2) of a native sequence PRO300 cDNA (nucleotides 45-1415), wherein the nucleotide sequence (SEQ ID NO:1) is designated herein as "UNQ263" and/or "DNA40625-1189". Also presented is the position of the initiator methionine residue.

Figure 2 shows the amino acid sequence (SEQ ID NO:3) derived from nucleotides 45-1415 of the nucleotide sequence shown in Figure 1.

Figures 3A-3B show an alignment of portions of the nucleotide sequences from a variety of expressed sequence tags as well as an intermediate consensus nucleotide sequence assembled therefrom designated "DNA35930".

Figures 4A through 4AH show an alignment of portions of the nucleotide sequences from a variety of expressed sequence tags as well as an extended consensus nucleotide sequence assembled therefrom using repeated cycles of BLAST and phrap, wherein the consensus sequence is herein designated <consen01>.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO300 polypeptide" and "PRO300" when used herein encompass native sequence PRO300 and PRO300 polypeptide variants (which are further defined herein). The PRO300 polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO300 polypeptide" comprises a polypeptide having the same amino acid sequence as a PRO300 polypeptide derived from nature. Such native sequence PRO300 polypeptide can be isolated from nature or can be produced by recombinant or

synthetic means. The term "native sequence PRO300 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of a PRO300 polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of a PRO300 polypeptide. In one embodiment of the invention, the native sequence PRO300 polypeptide is a mature or full-length native sequence PRO300 polypeptide comprising amino acids 1 to 457 of Figure 2 (SEQ ID NO:3). In another embodiment of the invention, the native sequence PRO300 polypeptide is an extracellular domain of the full-length PRO300 protein. Optionally, the PRO300 polypeptide is obtained or obtainable by expressing the polypeptide encoded by the cDNA insert of the vector DNA40625-1189 deposited on April 21, 1998 as ATCC _____.

60082707-0442200
"Percent (%) amino acid sequence identity" with respect to the PRO300 amino acid sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO300 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

"Percent (%) nucleic acid sequence identity" with respect to the UNQ263 (DNA40625-1189) sequence (SEQ ID NO:1) and nucleotides 45-1415 of the sequence shown in Figure 1 (SEQ ID NO:2) is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the UNQ263 (DNA40625-1189) sequence (SEQ ID NO:1) or nucleotides 45-1415 of the sequence shown in Figure 1 (SEQ ID NO:2), respectively, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the

00000797 - 042208

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO300 polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

5

50002707-042000

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO300 polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO300 antibody compositions with polyepitopic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Active" or "activity" for the purposes herein refers to form(s) of PRO300 which retain the biologic and/or immunologic activities of native or naturally-occurring PRO300 polypeptide.

The terms "treating", "treatment" and "therapy" as used herein refer to curative therapy, prophylactic therapy, and preventative therapy.

The term "mammal" as used herein refers to any mammal classified as a mammal, including humans, cows, horses, dogs and cats. In a preferred embodiment of the invention, the mammal is a human.

50082797-042296

II. Compositions and Methods of the Invention

A. Full-length PRO300 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO300. In particular, Applicants have identified and isolated cDNA encoding a PRO300 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO300 polypeptide have significant homology with the human Diff 33 protein. Accordingly, it is presently believed that PRO300 polypeptide disclosed in the present application is a newly identified member of the Diff 33 family.

B. Modifications of PRO300

Covalent modifications of PRO300 polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO300 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N-or C-terminal residues of a PRO300 polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO300 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO300 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO300 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one

or more carbohydrate moieties found in native sequence PRO300 polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO300 polypeptide.

Addition of glycosylation sites to PRO300 polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO300 polypeptide (for O-linked glycosylation sites). The PRO300 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO300 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO300 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO300 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo-and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO300 comprises linking the PRO300 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

PRO300 polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a PRO300 polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a PRO300 polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino-or carboxyl-terminus of the PRO300 polypeptide. The presence of such epitope-tagged forms of

50082797-042296

a PRO300 polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO300 polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a PRO300 polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

C. Preparation of PRO300

The description below relates primarily to production of PRO300 by culturing cells transformed or transfected with a vector containing at least nucleotides 45-1415 (SEQ ID NO:2) of the UNQ263 (DNA40625) nucleic acid (SEQ ID NO:1). It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO300 polypeptides. For instance, the PRO300 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of PRO300 polypeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce a full-length

PRO300 polypeptide.

1. Isolation of DNA Encoding PRO300

DNA encoding a PRO300 polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the PRO300 mRNA and to express it at a detectable level. Accordingly, human PRO300-encoding DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO300-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to a PRO300 polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO300 is to use PCR methodology [Sambrook et al., *supra*; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., *supra*.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTAR, and INHERIT which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the

60092797 " 042290

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO300 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., *supra*.

Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO_4 and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited

to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO300-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism.

Suitable host cells for the expression of glycosylated PRO300 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding the desired PRO300 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

The desired PRO300 polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO300-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO300-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene

provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, *Genetics*, **85**:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO300-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., *Nature*, **275**:615 (1978); Goeddel et al., *Nature*, **281**:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, *Nucleic Acids Res.*, **8**:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., *Proc. Natl. Acad. Sci. USA*, **80**:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO300 polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., *J. Biol. Chem.*, **255**:2073 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.*, **7**:149 (1968); Holland, *Biochemistry*, **17**:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO300 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding a PRO300 polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO300 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO300.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO300 polypeptides in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO300 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO300-encoding DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO300 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO300 polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO300 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO300 polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO300 polypeptide produced.

D. Uses for PRO300

Nucleotide sequences (or their complement) encoding PRO300 polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA.

862240 7622809

PRO300-encoding nucleic acid will also be useful for the preparation of PRO300 polypeptides by the recombinant techniques described herein.

The full-length nucleotide sequence SEQ ID NO:1 or the full-length native sequence PRO240 (SEQ ID NO:2) nucleotide sequence, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO300 gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of PRO300 or PRO300 from other species) which have a desired sequence identity to the PRO300 sequence disclosed in Figure 1 (SEQ ID NO:1). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the UNQ263 (DNA40625-1189) nucleotide sequence of SEQ ID NO:1 as shown in Figure 1 or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO300-encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO300 gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO300 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO300 sequences.

Nucleotide sequences encoding a PRO300 polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO300 polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO300 encode a protein which binds to another protein (example, where the PRO300 polypeptide functions as a receptor), the PRO300 polypeptide can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for

SECRET

SECRET

SECRET

060709Z JUL 84

4-

06-08-2797

2. Monoclonal Antibodies

The immunizing agent will typically include the PRO300 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused.

SECRET 7678200

immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against a PRO300 polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by

using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The anti-PRO300 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and

SECRET 042288

capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeven et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)).

SECRET 707 10/22/96

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a PRO300 polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming

a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

F. Uses for anti-PRO300 Antibodies

The anti-PRO300 antibodies of the present invention have various utilities. For example, anti-PRO300 antibodies may be used in diagnostic assays for PRO300 polypeptides, *e.g.*, detecting expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO300 antibodies also are useful for the affinity purification of PRO300 polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a PRO300 polypeptide are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO300 polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO300 polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO300 polypeptide from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1: Isolation of cDNA Clones Encoding Human PRO300

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA35930 (see Figures 3A and 3B). The DNA35930 consensus DNA sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended assembly sequence is indicated in numerous second alignment figures, as shown in Figures 4A-4AH, wherein the consensus sequence is therein designated <consen01>.

Based on the DNA35930 and <consen01> consensus sequences, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO300. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and

are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest by the *in vivo* cloning procedure using the probe oligonucleotide and one of the primer pairs.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 (35930.f1) 5'-GCCGCCTCATCTTCACGTTCTTCC-3' (SEQ ID NO:4)

forward PCR primer 2 (35930.f2) 5'-TCATCCAGCTGGTGCTGCTC-3' (SEQ ID NO:5)

forward PCR primer 3 (35930.f3) 5'-CTTCTTCCACTTCTGCCTGG-3' (SEQ ID NO:6)

forward PCR primer 4 (35930.f4) 5'-CCTGGGCAAAAATGCAAC-3' (SEQ ID NO:7)

reverse PCR primer 1 (35930.r1) 5'-CAGGAATGTAGAAGGCACCCACGG-3' (SEQ ID NO:8)

reverse PCR primer 2 (35930.r2) 5'-TGGCACAGATCTTCACCCACACGG-3' (SEQ ID NO:9)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35930 sequence which had the following nucleotide sequence

hybridization probe (35930.p1)

5'-TGTCATCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTG-3'
(SEQ ID NO:10)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO300 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue. The cDNA libraries used to isolate the cDNA clones were constructed by standard

[illegible]

The entire nucleotide sequence of UNQ263 (DNA40625-1189) is shown in Figure 1 (SEQ ID NO:1). Clone UNQ263 (DNA40625-1189) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 45-47 [Kozak et al., supra] and ending at the stop codon at nucleotide positions 1416-1418 (Figure 1). The predicted polypeptide precursor is 457 amino acids long (Figure 3). Clone UNQ263 (DNA40625-1189) has been deposited with ATCC and is assigned ATCC deposit no. _____.

EXAMPLE 2: Use of PRO300-encoding DNA as a hybridization probe

DNA comprising the coding sequence of full-length PRO300 (as shown in Figure 1, SEQ ID NO:1) is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO300) in human tissue cDNA libraries or human tissue genomic libraries.

28

performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO300 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO300 Polypeptides in *E. coli*

This example illustrates the preparation of unglycosylated forms of PRO300 polypeptides by recombinant expression in *E. coli*.

The DNA sequence encoding PRO300 (SEQ ID NO:3) is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO300 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO300 polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the polypeptide.

EXAMPLE 4: Expression of PRO300 Polypeptides in Mammalian Cells

This example illustrates preparation of glycosylated forms of PRO300 polypeptides by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO300-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO300-encoding DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO300.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO300 DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 µCi/ml ³⁵S-cysteine and 200 µCi/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO300 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO300-encoding DNA may be introduced into 293 cells transiently using the dextran sulfate method described by Sompayrac et al., *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 µg pRK5-PRO300 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture

medium, 5 µg/ml bovine insulin and 0.1 µg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO300 polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO300 polypeptide can be expressed in CHO cells. The pRK5-PRO300 vector can be transfected into CHO cells using known reagents such as CaPO_4 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S -methionine. After determining the presence of PRO300 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO300 polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO300 polypeptide may also be expressed in host CHO cells. The PRO300-encoding DNA may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO300-encoding DNA insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO300 polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

EXAMPLE 5: Expression of a PRO300 Polypeptide in Yeast

The following method describes recombinant expression of PRO300 polypeptides in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO300 polypeptide from the ADH2/GAPDH promoter. DNA encoding the PRO300 polypeptide of interest, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO300 polypeptide. For secretion, DNA encoding the PRO300 polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast

alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO300 polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO300 polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO300 polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 6: Expression of PRO300 Polypeptides in Baculovirus

The following method describes recombinant expression of PRO300 polypeptides in Baculovirus.

The PRO300-encoding DNA is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO300-encoding DNA or the desired portion of the PRO300-encoding DNA (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 to 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., *Baculovirus expression vectors: A laboratory Manual*, Oxford:Oxford University Press (1994).

Expressed poly-his tagged PRO300 polypeptide can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly,

SF9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+} -NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+} -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO300 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO300 polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 7: Preparation of Antibodies that Bind PRO300 Polypeptides

This example illustrates the preparation of monoclonal antibodies which can specifically bind to PRO300 polypeptides.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO300 polypeptide, fusion proteins containing a PRO300 polypeptide, and cells expressing recombinant PRO300 polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO300 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified

60052737-242200-001

in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO300 polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO300 polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO300 polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against a PRO300 polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO300 polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA40625-1189		April 21, 1998

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the

[illegible]

00000000000000000000000000000000

[illegible]

60082797-042209

WHAT IS CLAIMED IS:

1. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding a PRO300 polypeptide comprising the sequence of amino acid residues 1 to 457 of Figure 2 (SEQ ID NO:3), or (b) the complement of the DNA molecule of (a).
2. The nucleic acid of Claim 1 wherein said DNA comprises the nucleotide sequence of SEQ ID NO:1 or its complement.
3. The nucleic acid of Claim 1 wherein said DNA comprises nucleotides 45-1415 of the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2).
4. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit No. _____ (DNA40625-1189), or (b) the complement of the DNA molecule of (a).
5. The nucleic acid of Claim 4 which comprises a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit No. _____ (DNA40625-1189).
6. A vector comprising the nucleic acid of any one of Claims 1 to 5.
7. The vector of Claim 6 operably linked to control sequences recognized by a host cell transformed with the vector.
8. A host cell comprising the vector of Claim 6.
9. The host cell of Claim 8 wherein said cell is a CHO cell.
10. The host cell of Claim 8 wherein said cell is an *E. coli*.

[illegible]

13. Isolated native sequence PRO300 polypeptide comprising amino acid residues 1 to 457 of Figure 2 (SEQ ID NO:3).

15. A chimeric molecule comprising a PRO300 polypeptide fused to a heterologous amino acid sequence.

17. The chimeric molecule of Claim 15 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

18. An antibody which specifically binds to a PRO300 polypeptide.

37

Abstract of the Disclosure

The present invention is directed to novel polypeptides having homology to the human Diff 33 protein and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

SECRETARY OF STATE

[illegible]

002240 26228000

FIGURE 2

MGACLGACSLSCASCLCGSAPCILCCCPASRNSTVSRLIPTFFFLGLVLSIIMLSPG
VESQLYKLPWVCEGAGIPTVLQGHIDCGSLGAVYRMCFATAAFFFFFTLLMLCVS
SSRDPRAAIQNGFWFFKFLIIVGLTVGAFYIPDGSFTNIWFYFGVVGSEFLIILQVLII
DFAHSWNQRWLGKAEEDCSRAWYAGLFFFLIFLLSIAAVALMFMYYTEPSGCHEGKVF
ISLNLTFVCVCSIAAVALPKVQDAQPNQGLQASVITLYTMFVTWSALSSIPEQKCNPHLP
TQLGNETVVAGPEGYETQWDAPSI VGLIIFLLCTLFI SLRSSDHRQVNSLMQTECCPPM
LDATQOQQQVAAACEGRAFDNEQDGVITYSYSPFHFCVLVLSLHVMTLTNNWYKGETRKM
ISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNDFS

1971

FIGURE 3A

1226511	1	CAGCCGGGACCCCCGGGCTGCCATCCAGAAATGGGTTTGGTTCTTTAAGT
<DNA35930:	1	CAGCCGGGACCCCCGGGCTGCCATCCAGAAATGGGTTTGGTTCTTTAAGT
1226511	51	TCCTGATCTGGTGGGCTCACCGTGGGTGCGCTTCTACATTCCTGACGGC
1429622	1	ATCTGGTGGGCTCACCGTGGGTGCGCTTCTACATTCCTGACGGC
<DNA35930:	51	TCCTGATCTGGTGGGCTCACCGTGGGTGCGCTTCTACATTCCTGACGGC
1226511	101	TCCTTCACCAACATCTGGTTCTACTTCGGGTGCTGGGGCTCCTTCCTCTT
1429622	46	TCCTTCACCAACATCTGGTTCTACTTCGGGTGCTGGGGCTCCTTCCTCTT
1581204	1	TCGGTTCCTACTTCGGGTGCTGGGGCTCCTTCCTCTT
<DNA35930:	101	TCCTTCACCAACATCTGGTTCTACTTCGGGTGCTGGGGCTCCTTCCTCTT
1226511	151	CATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGGCACCTCTTGGAAAC
1429622	96	CATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGGCACCTCTTGGAAAC
1581204	38	CATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGGCACCTCTTGGAAAC
<DNA35930:	151	CATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGGCACCTCTTGGAAAC
1226511	201	ATCGCTGGCTGGGCAAGGCCGAGGAGTGCATTCCTCC-GTGCTGTGATCGC
1429622	146	AGCGGTGCTGGGCAAGGCCGAGGAGTGCATTCCTCC-GTGCTGTGATCGC
1581204	88	AGCGGTGGCTGGGCAAGGCCGAGGAGTGCATTCCTCC-GTGCTGTGNAACG
W49844	1	ATTGGCGAGGAGGGGATTCCTCAGTGCCTGTGATCGC
<DNA35930:	201	AGCGGTGGCTGGGCAAGGCCGAGGAGTGCATTCCTCAGTGCCTGTGATCGC
1226511	251	AGGCCTCTT
1429622	196	AGGCCTCTTCTTCTTCTACCTCTCTCTTCTACTTGTCTGT
1581204	138	AGGCCTCTTCTTCTTCTACCTCTCTCTTCTACTTGTGTGATCGCNGCCG
W49844	37	AGGCCTCTTCTTCTACCTCTCTCTTCTACTTGTGTGATCGCNGCCG
<DNA35930:	251	AGGCCTCTTCTTCTTCTACCTCTCTCTTCTACTTGTGTGATCGCNGCCG
1581204	188	TGGCGCTGATGTTCTATGTACTACAA
W49844	87	TGGCGCTGATGTTCTATGTACTACAAATGACCCAGCGGCTGCCACGAGGGC
<DNA35930:	301	TGGCGCTGATGTTCTATGTACTACAAATGACCCAGCGGCTGCCACGAGGGC

252240: 2628009

FIGURE 3B

W49844	137	AAGGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTGTGCGTGTCCATCGC
<DNA35930:	351	AAGGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTGTGCGTGTCCATCGC
W49844	187	TGCTGTCTTGCCTCCCAAGGTCCAGGACGCCAGCCCCAACTCGGGTCTGCTGC
<DNA35930:	401	TGCTGTCTTGCCTCCCAAGGTCCAGGACGCCAGCCCCAACTCGGGTCTGCTGC
W49844	237	AGGCCTCGGTTCATCACCTCTACACCATGTTTGTACCTGGTCAGCCCTA
<DNA35930:	451	AGGCCTCGGTTCATCACCTCTACACCATGTTTGTACCTGGTCAGCCCTA
W49844	287	TCCAAGTATCCCTGAACAAGAAATGCAACCCCATTTGCCAACCCAGCTG
<DNA35930:	501	TCCAAGTATCCCTGAACAAGAAATGCAACCCCATTTGCCAACCCAGCTG
W49844	337	GGCAACGAGACAGTTGTGGCAGGCCCGAGGGTATTGAGACCCAGTGGTGG
<DNA35930:	551	GGCAACGAGACAGTTGTGGCAGGCCCGAGGGTATTGAGACCCAGTGGTGG
W49844	387	GATGCCCCGAGCATTTGGGGCTCATCATCTTCTCCTGTGCACCTCTTC
<DNA35930:	601	GATGCCCCGAGCATTTGGGGCTCATCATCTTCTCCTGTGCACCTCTTC
W49844	437	ATCAGTCTGGGCTCCTCAGGCTCCACGGGACGGGGCTGCTGGAGAGAGCG
<DNA35930:	651	ATCAGTCTGGGCTCCTCAGGCTCCACGGGACGGGGCTGCTGGAGAGAGCG
W49844	487	GGGAACCTCCA
<DNA35930:	701	GGGAACCTCCA

FIGURE 4B

FIGURE 4B

1662702	246	CCC	
1929803	239	CCCTGGGTGTGTGANGAGGGGCCGGGATCCCCACCGTCTCT	
2520349	234	CCCTGGGTGTGT	
2487474	228	CCCTGGGTGTGTGA	
1223450	224	CCCTGGGTGTGTGAGGAGGGGCCGGGATCCCCACCGT	
774971	219	CCCTGGG	
2741948	217	CCCTGGGTGTGTGAGGAGGGGG	
1712126	217	CCC	
1843956	83	CCCTGGGTGTGTGAGGAGGGGCCGGGATCCCCACCGTCTCTGAGGGGCCA	
1434590	50	CCCTGGGTGTGTGAGGAGGGGCCGGGATCCCCACCGTCTCTGAGGGGCCA	
1434591	50	CCCTGGGTGTGTGAGGAGGGGCCGGGATCCCCACCGTCTCTGAGGGGCCA	
2121391	1	+++++ CAGGGCCA	
<consen01>	251	CCCTGGGTGTGTGAGGAGGGGCCGGGATCCCCACCGTCTCTGAGGGGCCA	
1843956	133	CATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCGCATGTGCT	
1434590	100	CATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCGCATGTGCT	
1434591	100	CATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCGCATGTGCT	
2121391	9	CATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCGCATGTGCT	
1705518	1	+++++ CGCATGTGCT	
<consen01>	301	CATCGACTGTGGCTCCCTGCTGGCTACCGCGTGTCTACCGCATGTGCT	
1843956	183	TGCGCAGGGCGGCTCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	
1434590	150	TGCGCAGGGCGGCTCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	
1434591	150	TGCGCAGGGCGGCTCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	
2121391	59	TGCGCAGGGCGGCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	
1705518	11	TGCGCAGGGCGGCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	
1226511	1	+++++ CTCTGCGTG	
<consen01>	351	TGCGCAGGGCGGCTCTCTCTCTTTTACCCGTCTATGCTCTGCGTG	

just

12

32	1581204	CCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGCGACTCCT	
20	2159726	CCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGCGACTCCT	
11	1815012	CCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGCGACTCCT	
9	HSU47700	CCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGCGACTCCT	
1	3233860	GCTGGTGTGCTCATCGACTTTGGCGACTCCT	
1	AA211364	CATCGACTTTGGCGACTCCT	
1	755491	ATCGACTTTGGCGACTCCT	
1	1712894	CTTTGGCGACTCCT	
1	616064	GNACTCCT	
<consen01>		+++..+...+++++...+++++...+++++...+++++...+..+++++	
550		CCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGGCGACTCCT	
209	1226511	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
195	DNA35930 .init	GGAAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCCAGTGCCTG	
140	1429622	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
82	1581204	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
70	2159726	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
61	1815012	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
59	HSU47700	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
33	3233860	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
21	AA211364	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCCAGTGCCTG	
20	755491	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
15	1712894	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
9	616064	GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCC - GTGCGCTG	
1	W49844	ATTGGGCAAGGAGGCGGAGTGCAGATTCCCAGTGCCTG	
1	493139	G	
<consen01>		+++++...+++++...+++++...+++++...+++++...+++++...++..+++++...	
600		GGAACACAGCGGTGGCTGGGCAAGCCGCGAGGAGTGCAGATTCCCAGTGCCTG	
258	1226511	GTACGCGAGCCTCTT	
245	DNA35930 .init	GTACGCGAGCCTCTTCTTCTTCACTCTCTCTTCTACTTGTCTGTCGATCG	
189	1429622	GTACGCGAGCCTCTTCTTCTTCACTCTCTCTTCTACTTGTCTGTCGT	
131	1581204	GNAOGCAGGCTCTTCTTCTTCACTCTCTCTTCTACTTGTCTGTCGATCG	
119	2159726	GTACGCGAGCCTCTTCTTCTTCACTCTCTCTTCTACTTGTCTGTCGATCG	
110	1815012	GTACGCGAGCCTCTTCTTCTTCACTCTCTCTTCTACTTGTCTGTCGATCG	

FIGURE 4G

[illegible]

25

FIGURE 4H

	101	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCCAACTC	
	100	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	97	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	94	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	91	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	57	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	55	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	48	GGTGCTGTGTTGAATGCTGTGGTGTAGCATATGTGGTTTTGCAGATT	
	19	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	12	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	5	CCTCGCGCTGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTN	
	1	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCAACTC	
	1	GGCAGGGGCTTTACCAGGGGTCCTCTTGCCAGGACGCCCAGCCAACTC	
	1+.....+.....+.....+.....+.....+..	
	795	GTGTCATCGCTGCTGCTGCCCAAGGTCAGGACGCCCAGCCCAACTC	
	440	GGGTC-T-GCTGCAGGCCCTGGTGCTATCATCCCTCTACACCATGTTGTGCAC	
	303	GGGTC-TGTTGAGGGCTCGGACATCACCCCTCTACACCATGTTGTGCAC	
	226	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	197	GGGTAAT-GCTGCAGGNCITAGGTGCTATCANCMNATACANCAATGNT	
	161	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	151	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	150	GGGTC-T-GCTNAGGCCCTNGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	147	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCT	
	144	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	141	GGGTC-T-GCTGCAGGNCITGGTGCTATCANCCCTCTACANCAATGTTGTGCAN	
	107	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	105	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	98	GGNGA-G-ANGSGCTCAGGTACGTCGTGTAAACCCGGGAAGCAGGTGATGG	
	62	GGGTC-T-GCTGAAGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	69	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	55	GGGTC-T-GCTGCAGGCCCTGGTGCTATNACCCNTTACACCATGTTGTGNAC	
	49	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
	50	GGGTC-T-GCTGCAGGCCCTGGTGCTATCACCCCTCTACACCATGTTGTGCAC	
<consen01>			
DDNA35930..init			
HSDU47700			
W49844			
Y493139			
I665446			
R72007			
R72471			
S356356			
R76679			
S820198			
S3320012			
I424153			
I1431207			
T2117449			
I1282436			
S50405			
I1312500			
I2377129			
I274786			

862240-26200003

FIGURE 41

274786	9	GGGTC-T-GCTGCAGGCTCGGTCAATNACCCCTCTACACCATGTTTAAAN
AA130551	1	TC-T-GCTGCAGGCTCGGTCAATCACCCTCTACACCATGTTTGTGAC
AA481928.RC	1	GCAGG-A-GGCACCATGGGCTGTGAGGCTGCCTCAGCTGAAGTCG
023897	1	CAGGCTCGNCCATACCCCTCTACACCATGTTTGTAAAC
1453117	1	CAGGCTCGGTCAATCACCCTCTACACCATGTTTGTGAC
1550018	1	ATCACCCCTCTACACCATGTTTGTGAC
1731292	1	CCTCTACACCATGTTTGTGAC
2417559	1	ACACCATGTTTGTGAC
AA410947.RC	1	GCTGAAGTCG
1754776	1	GTTTGTGAC
<consen01>	++.....
DNA35930.init	845	GGGTC T GCTGCAGGCTCGGTCAATCACCCTCTACACCATGTTTGTGAC
HSU47700	488	CTGGTCAGCCCTATCCAAAGTATCCCTGAAACAAGAAATGCAACCCCCCATTT
W49844	352	CTGGNCAGACCTATCCAN-TATCCCTGNAACA-GAAATGTAACCCCCCATTT
1665446	274	CTGGTCAGCCCTATCCAAAGTATCCCTGAAACAAGAAATGCAACCCCCCATTT
R72007	209	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAAC
R72471	199	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
876679	198	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
820198	192	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
3120012	189	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
1431207	155	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
2117449	153	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
1282436	146	AGGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
550405	117	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
1312500	110	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
2377129	103	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
274786	97	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
AA130551	98	CTGGTCAGCCCTATCCA-GTATCCCTGTAAGTATGGCCAGGCTCAAG
AA481928.RC	57	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
023897	46	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
1453117	49	CGGTTGCGCAGGAGGAGTGGGCTACCAAGGTCCACAGGTAGAGGAGCAG
1550018	39	CTGGNCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
	39	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT
	27	CTGGTCAGCCCTATCCA-GTATCCCTGNAACA-GAAATGCAACCCCCCATTT

FIGURE 4J

1731292	22	CTGGTCAGCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
2417559	17	CTGGTCAGCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
AA410947.RC	11	CGTGTCTGAGGAGGTGGCTACGAGGTCCACAGGTAGAGGAGCAG	
1754776	10	CTGGTCAGCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
2279337	1	GGTCAGCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
941498	1	GGTCAGCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
755392	1	GCCCTATCCA-GTATCCCTGAACA-GAATGCAACCCCAATTT	
548789	1	CCAGTATCCCTGNAACAGAAATGCAACCCCAATTT	
3039067	1	GTATCCCTGAACA-GAATGCAACCCCAATTT	
2376768	1	GCAACCCCAATTT	
1682539	1	GCAACCCCAATTT	
1804118	1	CGGCTCGAGT	
<consen01>	893	CTGGTCAGCCCTATCCAAGTATCCCTGAACAAGAAATGCAACCCCAATTT	
DNA35930.init	538	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
HSU47700	400	GGCAACNCAGNTGG-CAACGANA-AGTTGT-GG	
W49844	324	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
R72007	247	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
R72471	246	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
876679	240	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
820198	237	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
3120012	203	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
1424153	201	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
1431207	194	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
2117449	165	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
1282436	158	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
550405	151	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
1312500	145	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
2377129	147	GAGGCTGGCTCGT-TCTGGTCAATC-AG-TCTACTGTGGGCTG	
274786	105	TCAATNCAAGTGGNCAAGACANTTT-GG-CAGGCC-CAGGG-T	
AA130551	94	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
AA481928.RC	99	CCCTACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
023897	87	GNCAACCCAGTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	
1453117	87	GCCAAACCCAGCTGGG-CAACGAGACAGTTGT-GG-CAGGCC-CAGGG-T	

FIGURE 4K

```
<congen01>
```

274786	274786
AA130551	AA130551
AA481928.RC	AA481928.RC
023897	023897
1453117	1453117
1550018	1550018
1731292	1731292
2417559	2417559
AA410947.RC	AA410947.RC
1754776	1754776
2729337	2729337
941498	941498
755392	755392
548789	548789
3039067	3039067
2376768	2376768
1682539	1682539
1804118	1804118
2193380	2193380
3210913	3210913
2321125	2321125
2743359	2743359
1755244	1755244
1752950	1752950
3255760	3255760
1655567	1655567
AA477170.RC	AA477170.RC
H02520	H02520
2623580	2623580
732580	732580
1579688	1579688
1579888	1579888
794186	794186
1733505	1733505
Consens01>	Consens01>

FIGURE 4L

[illegible]

FIGURE 4M

DNA35930.init	629	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
W49844	415	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
R72471	339	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1282436	258	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2117449	251	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
550405	244	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1312500	238	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
274786	199	-NTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
AA130551	188	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
AA481928.RC	192	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
023897	180	TNTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1453117	180	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1550018	168	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1731292	163	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2417559	158	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
AA410947.RC	154	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1754776	151	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2279337	140	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
941498	140	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
755392	135	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
548789	129	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
3039067	125	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2376768	117	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1682539	107	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1804118	104	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2193380	89	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
3210913	90	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2321125	76	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
2743359	69	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1755244	69	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1752950	66	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
3255760	66	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
1655567	57	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
AA477170.RC	57	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG
H02520	45	-CTTCTCTCTGTGACACCTCTTTCATCATAGTCTGGCTCTCTCAGGCTCCACG

FIGURE 4N

[illegible]

FIGURE 40

115	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	1085	CAGGTGAA C AG CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG
106	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATG--TAG	294	ACGGCACAC
94	CAGG-TANA C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	285	ACGCCACACAGCAGCAG---CAGCAGCAGGTGGCAAGCCTGTGAGGGCCG
83	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	286	ACGCCACACAGCAGCAG---CAGCAG---GTGGCA-GCCTGTGAGGGCCG
76	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	249	ACGCCACACAGCAGCAG---CAGCAG---GTGGCA-GCCTGTGAGGGCCG
79	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	235	ACGCCACX
79	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	235	ACGCCACA
64	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	230	ACGNCAC
60	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	220	ACGCCACA
43	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	202	ACGCCACA
28	CAGGTGAA-C--AG-CCTGATGCAGACCGAGGAGTGCCTCCACCTATGCTAG	199	ACGCCACA
+ +++.....++++.....++..	184	ACGCCACA
		185	ACGCCACACAGCAGCAG---CAGCAG---GTGGCA-GCCTGTGAGGGCCG
		171	ACGCCACA
		164	ACGCCACA
		164	ACGCCACANNNNNNNN---NNNNNN---GTGGCA-GCCTGTGAGGGCCG
		164	ACGCCACANNNNNNNN---NNNNNN---GTGGCA-GCCTGTGAGGGCCG
		161	ACGCCACANNNNNNNN---NNNNNN---GTGGCA-GCCTGTGAGGGCCG
		151	ACGCCACACAGCAGCAG---CAGCAG---GTGGCA-GCCTGTGAGGGCCG
		139	ACGCCACACAGCAGCAG---CAGCAGCAGGTGGCA-GCCTGTGAGGGCCG
		129	ACGCCACANNNNNNNN---NNNNNN---GTGGCA-GCCTGTGAGGGCCG
33255760			
AAA47170.RC			
H023520			
H02623580			
H02732580			
H1579688			
H1794186			
H1733505			
H14503343			
H7242923			
<consen01>			
274786			
AAA130551			
AAA481928.RC			
AAA410947.RC			
2279337			
941498			
755392			
548789			
G039067			
1682539			
11804118			
2193380			
3210913			
2321125			
2743359			
11755244			
11752950			
33255760			
AAA47170.RC			
H02520			
H02623580			

732580	122	ACGNCACANNNNNNNNNN	---	NNNNNNNNNGTGGCA	-GCCTGTGAGGGCCCG
1579688	125	ACGCCACA			
1579888	125	ACGCCACA			
794186	110	ANGGCCACANNNNNNNNNN	---	NNNNNNNNNGTGGCA	-GCCTGTGAGGGCCCG
1733505	106	ACGCCACANNNNNNNNNN	---	NNNNNNNNNGTGGCA	-GCCTGTGAGGGCCCG
1450343	89	ACGCCACANNNNNNNNNN	---	NNNNNNNNNGTGGCA	-GCCTGTGAGGGCCCG
724923	74	ACGNCACANAGCAGNACGACG	AGCAG	CA	-GTGGCA-GCCTGTGAGGGCCCG
1310973	1				GTGGCA-GCCTGTGAGGGCCCG
836342	1				GTGGCA-GCCTGTGAGGGCCCG
836297	1				GTGGCA-GCCTGTGAGGGCCCG
1802601	1				GTGGCA-GCCTGTGAGGGCCCG
1529466	1				CA-GCCTGTGAGGGCCCG
		+.+.++++	+++++ ++++++ ++++++ ++++++
<conben01>	1131	ACGCCACACAGCAGCAG		CAGCAG	GTGGCA GCCTGTGAGGGCCCG
AA130551	332	GGCCTTTGACAA	CGAGCAGCAGCGCGGTCA	CACCTACAGCTAC	-TCCTTCTTTC
AA481928.RC	329	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CACCTACAGCTAC	-TCCTTCTTTC
AA410947.RC	292	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGGGAGGTGGCCCGGG	
3210913	228	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1755244	207	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	
1752950	207	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
3255760	207	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	
AA47170.RC	194	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
H02520	185	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
2623580	175	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
732580	168	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
794186	156	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1733505	152	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1450343	135	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
724923	120	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1310973	21	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
836342	21	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
836297	21	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1802601	21	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC
1529466	17	GGCCTTTGACAA	CGAGCAGCAGCAGCGCGGTCA	CCTACAGCTAC	-TCCTTCTTTC

```
<congen01>
```


FIGURE 4Q

1	238109	GGACGGCGTCACTACAGCTAC-TCC	TC
1	2528337	GGACGGCGTCACTACAGCTAC-TCC	TC
1174	<consen01>	GGACGGCGTCACTACAGCTAC-TCC	TC
381	AA130551	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
378	AA481928.RC	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
342	AA410947.RC	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
277	3210913	CGA-TTCTGCGCTGG-TGCTGGCGCTCACTG-CAGT	
243	AA477170.RC	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
234	HO2520	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
224	2623580	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
217	732580	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
205	794186	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
201	1733505	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
184	11450343	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
169	7724923	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
70	11310973	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
70	8836342	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
70	8836297	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
70	1802601	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
66	1529466	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
32	2238109	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
3	22528337	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	1228220	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	900944	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	900010	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	877907	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	2252522	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1	2272202	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	
1223	<consen01>	CAGTTCGCGCTGG-TGCTGGCGCTCACTG-CAGGTCATGATGACGCTCACC	

952210 46428009

FIGURE 4R

AA130551	429	AACTGGTACAAGNCCGGTGAG-AACCCGGAAGATGATGACACGTGG-ACC
AA481928.RC	426	AACTGGTACAAGCC
AA410947.RC	390	AACTGGTACAAGCC
AA477170.RC	291	AACTGGTACAAGCC
H02520	284	AACTGGTACAAGCCCGGTGAGGACCCGGAAGATGATCAGCACGTGGGACC
1450343	232	AACTGGTACAAGCCCGGTGAG-ACCCGGAA
724923	217	
1310973	118	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-A
836342	118	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
836297	118	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
1802601	118	AACTGGTACAAGCCCGGTGAG
1529466	114	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
238109	80	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
2528337	51	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
1228220	47	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
900944	27	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
900010	27	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
877907	11	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
2252522	5	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
2272202	5	AACTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
871317	1	CTGGTACAAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
1647217	1	AAGCCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
1731663	1	CCCGGTGAG-ACCCGGAAGATGATCAGCACGTGG-ACC
963070	1	GTGG-ACC
1688313	1	
R15713	1	
2173324	1	
<consen01>	1271	AACTGGTACAAGCCCGGTGAG ACCCGGAAGATGATCAGCACGTGG ACC

bioRxiv preprint doi: <https://doi.org/10.1101/262299>; this version posted May 1, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 4S

```
AA130551 477 GCCGTGT--GGTGAAGATTTGT-GCCAGCT-GGGCAAG-GGCTTGCTC
H02520 334 GCCGTGTGGGTGAAGATNT-GTTCAGCTNGGCA-G-GGGTGTTC
836342 166 NCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
836297 166 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
1529466 162 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
238109 128 GCCGTGT--GGTGAAGATCT-NT-GCCAGCT-GGGCA-G-GGCT-GCTC
2528337 99 GCCGTGT--GGTGAAGATCT-GT-GCCAACT-GGGCA-G-G-CT-GCTC
1228220 95 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
900944 75 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTN
900010 75 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
877907 59 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTG
2252522 53 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
2272202 53 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
871317 47 GCCGTGT--GGTGAAGATCT-GT-GCCANCT-GGGCA-G-GGCT-GCTC
1647217 40 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
1731663 37 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
963070 24 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
1688313 8 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
R15713 4 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGN-A-GAGGCT-GCTC
2173324 2 GCCGTGT--GGTGAAGATCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
H00901 1 TCT-GT-GCCAGCT-GGGCA-G-GGCT-GCTC
R67665 1 TCT-GT-GCCAGCT-GGG-A-G-GGCT-GCTC
2840414 1 GT-GCCAGCT-GGGCA-G-GGCT-GCTC
1315226 1 GT-GCCAGCT-GGGCA-G-GGCT-GCTC
1454892 1 GGT-GCTC
R73342 1 CT-GCTC
R73357 1 CT-GCTC
1981311 1 CTC
.>+++++ ++++++.*.* +++++.*. +.*. +.*. +.*.
1319 GCCGTGT GGTGAAGATCT GT GCCAGCT GGGCA G GGCT GCTC
<consen01>
```

AA130551
H02520
H36342
836297
1529466
238109
2528337
1228220
900944
900010
877907
2252522
2272202
871317
1647217
1731663
963070
1688313
L15713
2173324
H00901
R67665
2840414
1315226
1454892
R73342
R73357
1981311
1518815
AA211365
AA211365
1734407
2262038
2131632
<consen01

522	C
381	CT-NTTACCTGTTGGACCTGGGTAAAGGCCCAT
208	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTTCTTCGCGCAACCGCG-
208	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTNCTNCTGCGCAACCGNG-
204	CT-C
170	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG
140	CT-CT-ACCTGT-GGAACCTGGTAGCCCAATCCTCTCGCGCAACCG-G-
137	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
117	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTNCTGCGCAACCGNG-
117	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG
101	CT-CT-ACCTGT-GGACCCTGGTAGCCCANCTNCTGCGCAACCGCGCG
95	CTCT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
95	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
82	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
79	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
66	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTNCTNCTGCGCAACCGCG-
50	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
46	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCG-G-
44	CT-CT-ACCTGT-GGACCCTGGNAGCCCACTCCTCTCGCGCAACCGCG-
27	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
26	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGGN-
24	CT-CT-ACCTGT-NGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCN-
18	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
9	CT-CT-ACCTGT-GGACCCTGGTAGNCCCACTCCTCTCGCGCAACCGCG-
7	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
4	CT-CT-ACCTGT-GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG-
1	TGCGCAACCGCG-
1	CGGCG-
1	CGCG
1	GCGG
1	GCG-
++ .+ +++++ .++.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+.+	
CT CT ACCTGT GGACCCTGGTAGCCCACTCCTCTCGCGCAACCGCG	
[361]	

FIGURE 4T

FIGURE 4U

836342	254 A-C-TT-CAGCTGAGGCAGC-T-NACAGC-TGC-CATCTGGTGCTTCTGC
836297	254 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2528337	185 A- - -TT-CA-CTGAGGCAGC-T-CACAGCCTGC-NATCTGGTGCTTCTGC
1228220	183 A-C-TT-CAGCTGAGGCAGCCT-NACAGCCTGC-CATCTGGTGCTTCTGC
900944	163 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
900010	163 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
877907	148 A-C-TT-CAGCTGAGGCAGCCT-NACAGCCTGC-NATCTGGTGCTTCTGC
2252522	142 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2272202	141 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
871317	135 A-C-TT-NAGCTGAGGCAGCCT-NACAGCCTGC-CATCTGGTGCTTCTGC
1647217	128 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1731663	125 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
963070	112 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1688313	96 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
R15713	91 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2173324	90 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
H00901	73 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
R67665	71 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2840414	70 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1315226	64 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1454892	55 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
R73342	53 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
R73357	53 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1981311	50 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1518815	13 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
AA211365.RC	7 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
1734407	6 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2620380	5 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
2131632	4 A-C-TT-CAGCTGAGGCAGCCT-CACAGCCTGC-CATCTGGTGCTTCTGC
AA130587.RC	1 TT-CACAGNCTGC-CATCTGGTGCTTCTGC
R72869.RC	1 CTTC
237971	1 CTTC
<consen01>	1407 A C TT CAGCTGAGGCAGCCT CACAGCCTGC CATCTGGTGCTTCTGC

1862210" 204237005

FIGURE 4V

836342	297 C	CCA-NCTGGT-GCCT-NTNGNT-CGGT
836297	299	CCA-C-TGGTGGT-NTNGNT-CGGTGA-AAGCAAN-CN-GGNANCN
2528337	226	CCA-CCTGGT-GCCT-NTCGGT-NAA
1228220	228	NCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
900944	208	NCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
900010	208	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
877907	195	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
2252522	187	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
2272202	186	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
871317	180	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1647217	173	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1731663	170	CNA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1688313	157	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
963070	141	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
151815	137	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1315226	135	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1454892	118	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
733342	117	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
733357	115 C	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1981311	109	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
151815	100	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
AA211365.RC	99	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
1734407	98	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
2620380	58	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
2131632	52	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
AA130587.RC	51	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
R72869.RC	50	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
237971	49	CNA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
238502	29	CCAACNTGG-CTGT-CTNGGT-CGGTGA-CAGCCAA-CC-TNANCCN
R72008.RC	5	GNT-TTTCGT-GGCT-GGTTNA-AAGCN-AAATTGC-NC-CTTNCCTC
AA480065.RC	7	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TGCCCCCT
	1	CCA-CCTGGT-GCCT-CTNGGT-CGGTGA-CAGCCAA-CC-TGCCCCCT
	1	ATTG-CN-CCTTCNCC
	1	CCCCCT

662240 66228009

FIGURE 4W

<consen01>	1452	CCA	CTGGT	GCCT	CTGGCT	CGTGA	CAGCAA	CC	TGCCCT
2528337	269	-ANCAANAN	-AAN	-TAANCNAG	-NNNNAN	-ANGNA	-NNNNTN	-	NCN-NA
900944	251	-NCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
900010	251	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
877907	242	TCCNACAAGCA	ANTNAG	CAGNGCTGAG	-CCCC	-ACCCCT	-	GCC-NG	
2252522	230	-GGNCANAN	-CAA	-TCAGCCAG	-GCT				
2272202	229	-CCCCANAC	-CAA	-TCAGCCAG	-GCTGAN	-CNCCC	-ACCCCT	-	GCN-CC
871317	223	-TCCACAA	-CAA	-TNAGCCAG	-GGT				
1647217	216	-CCNA							
963070	199	-TCCANAA	-CAA	-TTAGNCAG	-GTTGGN	-CCCCA	-ACACCT	-	GCC-C-
1688313	184	-NNCCANAC	-CAA	-TNAGCCAG	-GCTGAN	-ANNCC	-ANCCCT	-	NNCNCC
R15713	180	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
2173324	178								
H00901	161	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
R67665	160	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
1315226	152	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAN	-CNCCC	-ACCCCT	-	GCC-CC
1454892	143	-CNCCANAC	-CAN	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCN-NC
R73342	142	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
R73357	141	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
1518815	101	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAN	-CNCCC	-ACCCCT	-	GCC-CC
AA211365.RC	96	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAN	-CNCCC	-ACCCCT	-	GCC-CC
1734407	94	-CCCCANAC	-CAA	-TCAGCCAG	-GCTGAN	-CNCCC	-ACCCCT	-	GCN-CC
2620380	93	-NCNCANAC	-CAN	-TNAGCCAG	-GCTGAN	-ANNCC	-ANCCCT	-	GNN-CC
2131632	91	-CCNNAC	-AA	-TNAGCCAG	-GCTGAN	-NGCCA	-CCCTGN	-	CCA-GT
AA130587.RC	72	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCC-CC
R72869.RC	48	-AAAAACNA	-AAT	-TAAGCCAA	-GNTGAG	-CCCC	-AACCTTTGCC	-	CC
237971	50	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
238502	44	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CC
R72008.RC	15	-CAAAACN	-AAA	-TAAGCCNA	-GCTGAG	-CCCC	-ACCCCT	-	GCC-CA
AA480065.RC	7	-CCCCACAC	-CAA	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCC-CC
1818182	1	CAC	-CAA	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCC-CC
2194762	1	CAC	-CAA	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCC-CC
1818702	1	CAC	-CAA	-TCAGCCAG	-GCTGAN	-CCCC	-ACCCCT	-	GCC-CC

FIGURE 4X

	CAC-CAA-TCAGCCAG-GCTGAN-CCCC-NCCCCT--GCC-CC	
AA010655	AAC10656.RC	1
R16077.RC	G--NCC-CA	1
R93337.RC	CT--GCC-CC	1
2053751	GCN-CC	1
H00902.RC	CC	1
<consen01> +. CCCCAC CAA TCAGCCAG GCTGAG CCCCC ACCCCT GCC CC	1495
2528337	310 A-NT	
900944	292 A-GT	
900010	292 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGG	
877907	288 A-GNT-T-CGAGG-AACTGNCNTTGA-G-CGGGCCCTT-G-TA-GTCG	
2272202	270 A-GCT-C-C-AGG-A-CCTGCCANT	
963070	239 A-GTTTC-AGG-A-NCTGCCCTT-GA-G-CGGGNCCTT-N-TA-GTTG	
1688313	226 A-GCT-C-C-AGG-A-CCTGC	
R15713	221 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
H00901	202 A-GCT-C-C-AGGA-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
R67665	201 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
1315226	193 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
1454892	185 A-GCT-N-C-AGG-A-CCTGCCCNT-GA-G-CGGGCCCTT-C-TA-GTNG	
R73342	183 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
R73357	182 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
1518815	142 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
NA11365.RC	136 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
1734407	135 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
2620380	135 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-NTGGGCCCTT-C-TA-GTCG	
2131632	130 C-AGG-A-N-TGN-C-TGAGCGGG-CT-T-TAGTGTAQT-N-CT-TANG	
AA130587.RC	112 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
R72869.RC	91 AAGNT-C-CCAGG-A-CCTGCCNT-GAAG-CGGGCCCTT-C-TAAGTCG	
237971	91 A-GCT-C-C-AGG-A-CCTGCCCTT-NA-G-CGGGCCCTT-C-TA-GTCG	
238502	85 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
R72008.RC	57 A-GNT-C-CCAGG-A-CCTGCCCTT-GAAG-CGGGCCCTT-CCTA-GTCG	
AA480065.RC	47 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	
1818182	37 A-GCT-C-C-AGG-A-CCTGCCCTT-GA-G-CGGGCCCTT-C-TA-GTCG	

FIGURE 4Y

2194762	37 A-GCT-C-C-AGG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
1818702	37 A-GCT-C-C-AGG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
AA010655	37 A-GCT-C-C-AGG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
AA010656.RC	15 A-GNT-C-C-AGG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
R16077.RC	7 A-GCT-C-CCAGG-A-CCTGCCCT-GA-N-CCGGGCCCTT-TAATTCG
R93337.RC	8 AAGCT-C-CCAGG-A-CCTGCCCT-GAAG-CCGGGCCCTT-C-TA-NTCG
2053751	6 A-NCT-C-C-ANG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
H00902.RC	3 A-GCT-C-CCAGG-A-CCTGCCCT-GA-AACCGGCCCTT-C-TA-ATCG
R72882.RC	1 TTC-C-C-AGN-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TAAGTCG
AA434486	1 C-C-AGG-A-CCTGCCCT-GA-G-CCGGGCCCTT-C-TA-GTCG
N94336.RC	1 GGCCTTTC-TAAGTCG
2758102	1 CCGGGGCCCTT-C-TA-GTCG
<consen01>	1536 A GCT C C AGG A CCTGCCCT GA G CCGGGGCCCT C TA GTCG
877907	329 TA-GTGGCCTT-CA--GGGTN-OG-AC
963070	279 TA-GTT
R15713	259 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-TCITGCAG
H00901	241 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
R67665	239 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-NCT-GCAG
1315226	231 TA-GTG-CCTT
1454892	223 TN-GTG-CNTT-NN--GGGTCC-GA
R73342	221 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-NCT-GCAG
R73357	220 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
1518815	180 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT
AA211365.RC	174 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
1734407	173 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
2620380	168 GC-CGA-GGAN-AT--AGGTCC-TG-AG-TTCCA-TCCCGG-CAC-ANCA
2131632	150 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
AA130587.RC	133 TA-GTG-CCTT-CA--GGGTCC-GAAGG-AGCAT-CAGGCT-CCT-GCAG
R72869.RC	129 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
237971	123 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
238502	98 TAAGTG-CCTT-CA--GGGTCC-GAAGG-AGCAT-NAAGCT-CCT-GCAG
R72008.RC	85 TA-GTG-CCTT-CA--GGGTCC-GA-GG-AGCAT-CAGGCT-CCT-GCAG
AA480065.RC	

1818102
2194762
1818702
AA010655
AA010656.RC
R16077.RC
R93337.RC
2053751
H00902.RC
R72882.RC
AA43486
N94336.RC
2758102
AA434390.RC
1284459
1979276
2325950
1554785
<<consen1>
R15713
H00901
R67665
R73342
R73357
AA211365.RC
1734407
26520380
2131632
AA130587.RC
R72869.RC
237971
R72008.RC

75	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
75	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
75	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
75	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
53	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAAGCT	CCT	GCAG	
48	TA	GTG	CTT	CA	-	GGGTCC	GAGG	AGCAT	CAAGCT	CCT	GNAG		
49	TA	GTG	CTT	NA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
44	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
43	TA	GTG	CTT	TCA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAN	
39	TA	GTG	CTT	CA	-	GGGTCC	GAGG	AGCAT	CAGGCT	CCT	GNAG		
35	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
16	TA	GTG	CTT	CA	AGGGTCC	GA	GGGAGCAT	CAAGNT	TCCTG	GCAG			
17	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
1	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
1						G	GTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG
1						G	CTT	CA	-	GGGTCC	GA	GG	AGCAT
1										CC	GA	GG	AGCAT
1													GCAG
1574	TA	GTG	CTT	CA	-	GGGTCC	GA	GG	AGCAT	CAGGCT	CCT	GCAG	
299	AGCC	AT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	TGCT
280	AGCCCAT	CCCCC	-	G	CC	-	ACANCCA	CACGTTGGAGNT	TGCTCTTT				
278	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	TGCTCTT	-
260	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT
259	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT
212	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT
212	AGCCCAT	-	CCCCC	-	G	CC	-	ACAC					
212	AGCCCAT	-	TNCNC	-	G	AN	-	ACACCC	CACGCTGG	-	AGCT	-	GNNTCAT
207	ACGCTGAN	-	TNCTTT	-	C	TT	-	CCTCTCTGTNCAAT	-	ANAT	-	T	
189	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT
173	AGCCCAT	-	CCCCC	-	G	CC	-	ANACCCAN	CACGCTGG	-	AGCT	-	GCCTCTT
168	AGCC	AT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-
162	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT
139	AGCCCAT	-	CCCCC	-	G	CC	-	ACACCC	CACGCTGG	-	AGCT	-	GCCTCTT

FIGURE 4AA

AA480065.RC	124	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
1818182	114	AGCCCCAT-NCCCC-G-CC-ANACCCACACGGTGG-AGCT-GCCTCTT-
2194762	114	AGCCCCAT-CCCCC-G-CN-ACACCCACACGGTGG-AGCT-GCCTCTT-
1818702	114	AGCCCCAT-CNCCC-G-CC-ACNCCACACGGTGG-AGCT-GCCTCTT-
AA010655	114	AGCCC-AT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
AA010656.RC	92	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
R16077.RC	90	AGCCCCAT-CCCCC-GNCC-AAACCCANANGTGG-AGCT-GCCTCTT-
R93337.RC	89	AGCCCCAT-CCCCC-G-CC-AAACCCANACGGTGG-AGCT-GCCTCTT-
2053751	83	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
H00902.RC	84	AGCCCCAT-NCCCC-G-CC-AAACCAA-ANGTGG-AGCT-GCCTCTT-
R72882.RC	80	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
AA434486	74	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
N94336.RC	61	AGCCCCAT-CCCCCG-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
2758102	56	ANCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
AA434390.RC	40	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
1284649	36	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
1979276	27	ANCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
2325950	25	ANCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
1554785	5	AGCCCCAT-CCCCC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
R66066.RC	1	CCCCAT-TCCCC-GNCC-AAACCCANANGTGG-AGCT-GCCTCTT-
807796	1	AT-NCCNCC-G-CC-ANANCCACACGGTGG-AGCT-GCCTCTT-
630305	1	CC-G-CC-ACACCCACACGGTGG-AGCT-GCCTCTT-
3277068	1	GG-AGCT-GCCTCTT-
T40555.RC	1	CT-GCCTCTT-
<consen01>	1613	+.....+ :..+.....+.....+.....+.....+.....+.....+ AGCCCCAT CCCCC G CC ACACCCACACGGTGG AGCT GCCTCTT

862240 4628009

FIGURE 4AB

H00901	327	CCTTCCCT-TCIN-CCGTGCGCATTAT-TTCAGGATTTTGGGAT
R67665	322	CCTT-CGNTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R73342	302	CCTT-CGCTT-CCTGCGCTTT-GCCCATANT-CAGCATTTTGGATGAAA
R73357	303	CCTT-CGCTTCCCTC-CCTGTT-GCCCATAAITCAGCATCTCGGATGAAA
AA211365.RC	255	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
2620380	255	CCTT-NCCTT-CCTC-CCTGTT-GCCCATATA
AA130587.RC	232	CCTT-CGCTT-CGNTN-CGNTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R72869.RC	216	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
237971	210	CCTT-CGCTT-CCTC-CCTGTT-GC
238502	204	CCTT-CGCTT-CCTC-CCTGTT-GC
R72008.RC	183	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
AA480065.RC	167	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
1818182	157	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
2194762	157	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
1818702	157	CCTT-CGCTT-CCTC-CCTGTT-GCCCATATA-T-CAGCATCTCGGATGAAA
AA010655	156	CCTT-CGCTT-CGNTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
AA010656.RC	135	CCTT-CGCTT-CGNTN-CGNTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R16077.RC	134	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R93337.RC	132	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
2053751	126	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
H00902.RC	124	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R72882.RC	117	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
AA434486	105	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
N94336.RC	99	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
2758102	83	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
AA434390.RC	79	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
1284649	70	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
1979276	68	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
2325950	48	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
1554785	43	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
R66066.RC	38	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
807796	32	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
630305	14	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
3277068	10	CCTT-CGCTT-CCTC-CCTGTT-GCCCATACT-CAGCATCTCGGATGAAA
T40555.RC		

FIGURE 4AD

807796	83	GGGNT-NCTTGT-CCT-CAGGCT-NCACGGG-AGCGGGGGC-T-GCT-
630305	77	GGGCT-CCTTGT-CCT-CAGGCT-CCACGGG-AGCGGG-C-T-GCT-
3277068	59	GGNT-CCTTGT-CCT-CAGGCT-CCACGGG-AGCNGGG-C-T-GCT-
T40555.RC	55	GGCT-CCTTGT-CCT-CAGGCT-CCACGGG-AGCGGG-C-T-GCT-
756679	33	GGNT-NCTTGT-CCT-CAGGCT-TCACGGG-AGCGGG-C-T-NCT-
AA047596.RC	29	GGCT-CCTTGT-CNT-CNAGNT-CCACGGG-AGCGGG-C-T-GCT-
3142901	23	GGCT-CCTTGT-CCT-CAGGCT-CCACGGG-AGCGGG-C-T-GCT-
762960	12	GGCT-CCTTGT-CCT-CAGGCT-CCACGGG-AGCGGG-C-T-GCT-
272902	1	CCCTGT-CCT-CAGGCT-CCACGGG-AGCNGGG-C-T-GCT-
1978442	1	CCCTGT-CCT-CAGGCT-CCACGGG-AGCNGGG-C-T-GCT-
2306960	1	CCCTGT-CCT-CAGGCT-CCACGGG-AGCNGGG-C-T-GCT-
1926134	1	GC-A-A-CCG-GCT-
<consen01>	1701	GGGCT CCTTGT CCT CAGGCT CCACGGG AGCGGG C T GCT
H00901	414	NC-GGGGGTNTT--GG-G-AGAAGAGC-GG-GGGGAATTTCCACAA
R67665	409	GGGAGAGCNGGG--GA-A-TT
R73342	397	GG-AGAGAGCGGGGAA-T-TCCACCAACA-TTGGGGNATTCGGNA
R73357	394	GGGAGAGAGCGGGGAA-T-TNCCACCAACA-TTGGGGCAAT
AA211365.RC	340	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
AA130587.RC	317	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
R72869.RC	300	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
R72008.RC	267	GG-AGAGAGCGGG--AA-N-CTCNACCA-NA-GTGGGGCATCCGGNACT
AA480065.RC	252	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
1818182	242	GG-AGAGAGCGGG--AA--TCCCAAC
2194762	242	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCG
1818702	239	GG-AAAAA-CGGGG--AA-N-TCCNCCAA-AG-GGGGNTCGGNATTA
AA010655	240	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGG-ACT
AA010656.RC	220	GG-AGAGAGCGGG--AA-C-TCC-ACCA-CA-GTGGGGCATCCGGCACT
R16077.RC	219	GG-AGAGAGCGGG--ANAC-TNC-ACCA-CA-GTGGGGCATCCGGCACT
R93337.RC	216	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
2053751	211	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
H00902.RC	210	GG-AGAGAGCGGG--NA-C-TCCACCA-NA-GTGGGGCATCCGGNACT
R72882.RC	208	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
AA434486	201	GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT

662240 66228009

FIGURE 4AE

```
N94336.RC
2758102
AA434390.RC
1284649
1979276
2325950
1554785
R66066.RC
807796
630305
3277068
T40555.RC
756679
AA047596.RC
3142901
762960
2729902
1978442
2306960
1926134
1737558
1735126
1929940
AA283209.RC

189 GG-AGAGAGCGGG--AA-C-TTCCACCA-CA-GTGGGGCATCCGGCACT
184 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
168 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
164 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
155 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
153 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
133 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
127 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
124 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
117 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
99 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
95 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
73 NG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
69 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
63 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
52 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
36 GG-AGAGAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
36 GG-AGANAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
8 CG-ACCTCAGCTGA-NG-C-AGCCTCAC-AG-CCTGCCATCTGTGCT
4 GG-AGANAGCGGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
1 GGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
1 GGG--AA-C-TCCACCA-CA-GTGGGGCATCCGGCACT
1 C-TCCACCA-CA-GTGGGGCATCCGGCACT
1 CCCACCA-CA-GTGGGGCATCCGGCACT
1741 GG-AGAGAGCGGG AA C TCCACCA CA GTGGGGCATCCGGCACT
<consen01>
H00901
R73342
AA211365.RC
AA130587.RC
R72869.RC
R72008.RC
AA480065.RC
AA010655
457 ATTGGGGGCAATTCGG
443 ATTGAAGGCTTGTCTTCT
383 GAAGCCCTGGTCTCTGT-C-A-CGT-CNCCCAGGGACCTTCCCGC
360 GAAGCCCTGGTCTCTGT-C-A-NTT-CCCCAGGG-ACCCTGCCCCC
344 GAAGCCCTGGTCTCTGT-C-ANGCT-CCOC-AGGGACCTTCCCGC
310 GAAGCCCTGNTGNCCTGTNC-A-CCG-CGCCCAGGG-ACCCTGCCCCC
295 GAAGCCCTGGTCTCTGT-C-A-CGT-CCCCAGGGACCTTCCCGC
283 GAAGCCCTGGTCTCTGT-C-A-CGT-CCCC-ANGGGACC
```

四、

FIGURE 4A-F

1784

```
<congen01>
```

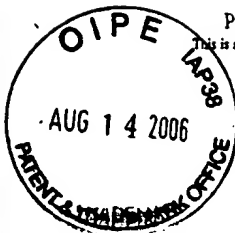

AA211365.RC
AA130587.RC
R72869.RC
R72008.RC
AA480065.RC
AA010656.RC
R16077.RC
R33337.RC
H00902.RC
R72882.RC
AA34486
N94336.RC
AA434390.RC
2325950
R65066.RC
G07796
330305
3277068
4740555.RC
756679
AA047596.RC
3142901
G62960
2372902
1978442
2306960
926134
7737558
1735126
9229940
AA283209.RC
1126005
consen01>

[illegible]

RefSeq: 2628009

FIGURE 4AH

AA211365.RC	479 AGCCAGTGGGTGTAA	
AA130587.RC	455 AGCCAGTGGGTGTACCATGAAAA	
R72869.RC	440 AGNNAGTNCNTGTAAANAAN	
R72008.RC	406 NNTTNAAGTCCGTGNAATTT	
AA480065.RC	391 AGCCAGTGGGTGTACCATGAAAA	
AA010656.RC	358 AGCCAGTGGGTGTACCN	
R16077.RC	358 AGCCAGTGGGTGTACCAAAAAA	
R93337.RC	355 AGCCAGTGGGTGTTCANATNA	
H00902.RC	349 AGCCAGT	
R72882.RC	347 NG	
AA434486	340 AGCCAGTGGGTGTACCA	
N94336.RC	327 AGCCAGTGGGTGTAAAAA	
AA434390.RC	308 AGCCAGTGGGTGTACCAAAAAA	
R66066.RC	266 AGCCAGTNCGTGTNCCAAAAANT	
630305	254 AGNCANTGCGT	
3277068	238 AGCCAGTGGGTGT	
T40555.RC	233 AGCCAGTGGGTGT	
756679	213 TTGTTAATAAACCACCAAGTGGCGTTGTAAAAAANNAANN	
AA047596.RC	209 AGCCAGTGCNTGTACC	
3142901	202 AGCCAGTGGGTGTAAAA	
762960	191 AGCCAGTGGGTGT	
2729902	175 AGCCAGTGGGTGT	
1978442	175 AGCCAGTGGGTGTAA	
2306960	147 AGCCAGTGGGTGTACCATG	
1926134	143 AGCCAGTGGGTGTAC	
1737558	131 AGCCAGTGGGTGTAC	
1929940	125 AGCCAGTGGGTGTAAAA	
AA283209.RC	123 AGCCAGTGGGTGTACCATGAAAA	
2126005	64 AGCCAGTGGGANTACCA	
<consen01>eeeeeeeeeeeeeeeeeeee
756679	1880 AGCCAGTGGGTGTACCATGAAAAAGTGGCGTTGTAAAAAANNAANN	
<consen01>	263 ANA	
	eee	
	1930 ANA	



PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

"EXPRESS MAIL" MAILING LABEL

Number: EM 178 612 241 115 Date of Deposit: April 29, 1998

I hereby certify that this paper or file is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Typed or Printed Name: Kron, Donald

Signed: Kron

A/PROV

Docket	P-66091/WH/D/DAV
--------	------------------

INVENTOR(s)/APPLICANT(s)					
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)		
WOOD	WILLIAM	I.	San Mateo, California		
TITLE OF THE INVENTION (280 characters max)					
NOVEL MULTI-TRANSMEMBRANE POLYPEPTIDES HAVING SEQUENCE IDENTITY WITH MEMBRANE REGULATOR PROTEINS AND NUCLEIC ACIDS					
CORRESPONDENCE ADDRESS					
Walter H. Dreger FLEHR HOHBACH TEST ALBRITTON & HERBERT LLP, 4 Embarcadero Center, Suite 3400, San Francisco					
STATE	CA	ZIP CODE	94111	COUNTRY	US
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification	Number of Pages	41	<input type="checkbox"/>	Small Entity Statement
<input checked="" type="checkbox"/>	Drawings	Number of Sheets	6	<input type="checkbox"/>	Other (specify):
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the Provisional filing fees			PROVISIONAL FILING FEE AMOUNT (\$)	\$150
<input checked="" type="checkbox"/>	The Commissioner is hereby authorized to charge any underpayment or overpayment to Deposit Account Number: 06-1300 (Order No. P-66091/WH/D/DAV)				

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒

No.

☐

Yes, the name of the U.S. Government Agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE:

Donald Kron 39054

Date:

April 29, 1998

TYPED or PRINTED NAME: Walter H. Dreger

REGISTRATION NO.
(if appropriate)

24,190

☐

Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take .2 hours to complete. Times will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Assistant Commissioner for Patents, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

583498



PATENT
Attorney Docket No.: P-66091/WH/DAV
(Genentech Docket No.: PR1335)

NOVEL MULTI-TRANSMEMBRANE POLYPEPTIDES HAVING SEQUENCE
IDENTITY WITH MEMBRANE REGULATOR PROTEINS AND NUCLEIC ACIDS
ENCODING THE SAME

FIELD OF THE INVENTION

5 The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel multi-transmembrane polypeptides having sequence identity with membrane regulator proteins, designated herein as "PRO218" polypeptides.

BACKGROUND OF THE INVENTION

10 Membrane-bound proteins and receptors can play an important role in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which
15 are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of
20 signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

25 Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interaction. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

Efforts are being undertaken by both industry and academia to identify new, native

membrane bound proteins, particularly those having sequence identity with membrane regulator proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins.

SUMMARY OF THE INVENTION

5 Applicants have identified a cDNA clone that encodes a novel multi-transmembrane protein having sequence identity with membrane regulator proteins, wherein the polypeptide is designated in the present application as "PRO218".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO218 polypeptide. In one aspect, the isolated nucleic acid
10 comprises DNA encoding the PRO218 polypeptide having amino acid residues 1 through 455 of Figure 2 (SEQ ID NO:2), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. The isolated nucleic acid sequence may comprise the cDNA insert of the vector deposited on 28 April 1998 with the ATCC as DNA30867-1335 which includes the nucleotide
15 sequence encoding PRO218.

In another embodiment, the invention provides a vector comprising DNA encoding a PRO218 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO218 polypeptides is further provided and comprises culturing host cells under conditions suitable
20 for expression of PRO218 and recovering PRO218 from the cell culture.

In another embodiment, the invention provides isolated PRO218 polypeptide. In particular, the invention provides isolated native sequence PRO218 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 455 of Figure 2 (SEQ ID NO:2). Optionally, the PRO218 polypeptide is obtained or is obtainable by
25 expressing the polypeptide encoded by the cDNA insert of the vector deposited on 28 April 1998 with the ATCC as DNA30867-1335.

In another embodiment, the invention provides chimeric molecules comprising a PRO218 polypeptide fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO218 polypeptide fused to an epitope tag sequence
30 or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to a PRO218 polypeptide. Optionally, the antibody is a monoclonal antibody.

In another embodiment, the invention provides expressed sequence tags (ESTs)

comprising the nucleotide sequence of SEQ ID NO:3 or 4.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show a nucleotide sequence (SEQ ID NO:1) containing the nucleotide sequence of a native sequence PRO218 cDNA (nucleotides 150-1514), wherein the nucleotide sequence (SEQ ID NO:1) is a clone designated herein as "UNQ192" and/or "DNA30867-1335". Also presented is the position of the initiator methionine residue, circled, and the stop codon, circled. The putative transmembrane domains of the protein are encoded by nucleotides beginning at nucleotides 258, 390, 597, 1011, 1161, 1272, 1422 in the figure. The complementary sequence and deduced amino acid sequence are also shown.

Figures 2A and 2B show amino acid sequences. Figure 2A shows the amino acid sequence (SEQ ID NO:2) derived from nucleotides 150-1514 of the nucleotide sequence shown in Figures 1A and 1B. Also shown in Figure 2B are a signal peptide, potential transmembrane domains, N-glycosylation sites, and a eukaryotic cobalamin-binding protein.

Figure 3 shows two novel ESTs, SEQ ID NOS:3 and 4.

Figure 4 shows from DNA or consensus DNA formed on the basis of SEQ ID NOS:3 and 4.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "PRO218 polypeptide" and "PRO218" when used herein encompass native sequence PRO218 and PRO218 polypeptide variants (which are further defined herein). The PRO218 polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO218 polypeptide" comprises a polypeptide having the same amino acid sequence as a PRO218 polypeptide derived from nature. Such native sequence PRO218 polypeptide can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO218 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of a PRO218 polypeptide (e.g., soluble forms containing for instance, an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of a PRO218 polypeptide. In one embodiment of the invention, the native sequence PRO218 polypeptide is a full-length or mature native sequence PRO218 polypeptide comprising amino acids 1 or 24 through 455 of SEQ ID NO:2. In another embodiment of the invention, the native

866240-5643000

sequence PRO218 polypeptide is an extracellular domain of the full-length PRO218 protein. Optionally, the PRO218 polypeptide is obtained or obtainable by expressing the polypeptide encoded by the cDNA insert of the vector DNA30867-1335 deposited on 28 April 1998 with the ATCC.

The "PRO218 extracellular domain" or "PRO218 ECD" refers to a form of the PRO218 polypeptide which is essentially free of the transmembrane and cytoplasmic domains of the PRO218 polypeptide. Ordinarily, PRO218 ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domain identified for the PRO218 polypeptide of the present invention is identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified.

"PRO218 variant" means an active PRO218 polypeptide as defined below having at least about 80% amino acid sequence identity with the PRO218 polypeptide having the deduced amino acid sequence shown in Figure 2A (SEQ ID NO:2) for a full-length or mature native sequence PRO218 polypeptide. Such PRO218 polypeptide variants include, for instance, PRO218 polypeptides wherein one or more amino acid residues are added, substituted or deleted, at the N- or C-terminus or within the sequence of Figure 2 (SEQ ID NO:2). Ordinarily, a PRO218 polypeptide variant will have at least about 80% amino acid sequence identity, preferably at least about 85% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity and even more preferably at least about 95% amino acid sequence identity with the amino acid sequence of Figure 2 (SEQ ID NO:2), with or without the signal peptide. The variants provided herein exclude the polypeptides and nucleic acids described herein which the novel polypeptides have identity with and which are already known in the art.

"Percent (%) amino acid sequence identity" with respect to the PRO218 amino acid sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO218 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software.

Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Similarity can be determined using the same programs.

"Percent (%) nucleic acid sequence identity" with respect to the PRO218 sequence identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO218 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Similarity can be determined using the same programs.

The term "epitope tagged" where used herein refers to a chimeric polypeptide comprising a PRO218 polypeptide, or domain sequence thereof, fused to a "tag polypeptide".

15 The tag polypeptide has enough residues to provide an epitope against which an antibody may be made, or which can be identified by some other agent, yet is short enough such that it does not interfere with the activity of the PRO218 polypeptide. The tag polypeptide preferably is also fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between

20 about 8 to about 50 amino acid residues (preferably, between about 10 to about 20 residues).

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO218 polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO218 polypeptide-encoding nucleic acid molecule is a nucleic acid

molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO218 polypeptide-encoding nucleic acid. An isolated PRO218 polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated PRO218 polypeptide-encoding nucleic acid molecules therefore are distinguished from the PRO218 polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated PRO218 polypeptide-encoding nucleic acid molecule includes PRO218 polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express PRO218 polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO218 polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO218 antibody compositions with polypeptidic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Active" or "activity" for the purposes herein refers to form(s) of PRO218 which retain

6006343X.042996

the biologic and/or immunologic activities of native or naturally-occurring PRO218 polypeptide.

The terms "treating", "treatment" and "therapy" as used herein refer to curative therapy, prophylactic therapy, and preventative therapy.

The term "mammal" as used herein refers to any mammal classified as a mammal, including humans, cows, horses, dogs and cats. In a preferred embodiment of the invention, the mammal is a human.

II. Compositions and Methods of the Invention

A. Full-length PRO218 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO218. In particular, Applicants have identified and isolated cDNA encoding a PRO218 polypeptide, as disclosed in further detail in the Examples below. The PRO218-encoding clone was isolated from a human fetal kidney library. To Applicants present knowledge, the UNQ192 (DNA30867-1335) nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no significant sequence identities to any known proteins were revealed. Some sequence identity was found with membrane regulator proteins, indicating that PRO218 may function as a membrane regulator.

B. PRO218 Variants

In addition to the full-length native sequence PRO218 polypeptide described herein, it is contemplated that PRO218 variants can be prepared. PRO218 variants can be prepared by introducing appropriate nucleotide changes into the PRO218-encoding DNA, or by synthesis of the desired PRO218 polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO218 polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO218 or in various domains of the PRO218 polypeptide described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO218 polypeptide that results in a change in the amino acid sequence of the PRO218 polypeptide as compared with the native sequence PRO218. Optionally the

variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO218 polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO218 polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in any of the *in vitro* assays described in the Examples below.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 15 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO218-encoding variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

C. Modifications of PRO218

Covalent modifications of PRO218 polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO218 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of a PRO218 polypeptide.

Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO218 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO218 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutamyl and asparagyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO218 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. 15 "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO218 polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO218 polypeptide.

20 Addition of glycosylation sites to PRO218 polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO218 polypeptide (for O-linked glycosylation sites). The PRO218 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO218 polypeptide at preselected bases such that codons are generated
25 that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO218 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Apelin and Wriston, *CRC Crit. Rev. Biochem.*, pp. 259-306 (1981).

30 Removal of carbohydrate moieties present on the PRO218 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch.

866240-5545809

Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO218 comprises linking the PRO218 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

PRO218 polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a PRO218 polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a PRO218 polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO218 polypeptide. The presence of such epitope-tagged forms of a PRO218 polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO218 polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a PRO218 polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

The PRO218 polypeptide of the present invention may also be modified in a way to form a chimeric molecule comprising a PRO218 polypeptide fused to a leucine zipper.

Various leucine zipper polypeptides have been described in the art. See, e.g., Landschulz et al., Science 240:1759 (1988); WO 94/10308; Hoppe et al., FEBS Letters 344:1991 (1994); Maniatis et al., Nature 341:24 (1989). It is believed that use of a leucine zipper fused to a PRO218 polypeptide may be desirable to assist in dimerizing or trimerizing soluble PRO218 polypeptide in solution. Those skilled in the art will appreciate that the leucine zipper may be fused at either the N- or C-terminal end of the PRO218 molecule.

D. Preparation of PRO218

The description below relates primarily to production of PRO218 by culturing cells transformed or transfected with a vector containing PRO218 polypeptide encoding nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO218 polypeptides. For instance, the PRO218 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of PRO218 polypeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce a full-length PRO218 polypeptide.

1. Isolation of DNA Encoding PRO218

DNA encoding a PRO218 polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the PRO218 mRNA and to express it at a detectable level. Accordingly, human PRO218-encoding DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO218-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to a PRO218 polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO218 is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring

Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Included in this invention are nucleic acids having at least about 70% to about 75% nucleic acid sequence identity with the nucleic acid sequence shown in Figure 1A and 1B (SEQ ID NO:1), with or without the untranslated portions and with or without the region encoding the signal peptide. Ordinarily, in another embodiment, the nucleic acids will have at least about 80% nucleic acid sequence identity, preferably at least about 85% nucleic acid sequence identity, more preferably at least about 90% nucleic acid sequence identity and even more preferably at least about 95% nucleic acid sequence identity with the nucleic acid sequence of Figure 1A and 1B (SEQ ID NO:1), with or without the untranslated portions and with or without the region encoding the signal peptide.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTar, and INHERIT which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO218 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature,

pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

- Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO_4 and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 21:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

- Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO218-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism.

- Suitable host cells for the expression of glycosylated PRO218 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include

monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding the desired PRO218 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art.

Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The desired PRO218 polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO218-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors
5 in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from
10 complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO218-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub
15 et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 (Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)). The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

20 Expression and cloning vectors usually contain a promoter operably linked to the PRO218-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*)
25 promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO218 polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the
30 promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-

50083495-042998

phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with
5 nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO218 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK
10 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding a PRO218 polypeptide by higher eukaryotes may be
15 increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic
20 cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO218 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal,
25 human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as
30 polyadenylated fragments in the untranslated portion of the mRNA encoding PRO218.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO218 polypeptides in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO218 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO218-encoding DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO218 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO218 polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO218 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO218 polypeptide. Various methods of protein purification

may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO218 polypeptide produced.

5

E. Uses for PRO218

Nucleotide sequences (or their complement) encoding PRO218 polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA.

- 10 PRO218-encoding nucleic acid will also be useful for the preparation of PRO218 polypeptides by the recombinant techniques described herein.

The full-length DNA30867-1335 nucleotide sequence (SEQ ID NO:1) or the full-length native sequence PRO218 nucleotide sequence, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO218 gene or to isolate
15 still other genes (for instance, those encoding naturally-occurring variants of PRO218 or PRO218 from other species) which have a desired sequence identity to the PRO218 nucleotide sequence disclosed in Figure 1A and 1B (SEQ ID NO:1). Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the UNQ192 (DNA30867-1335) nucleotide sequence of SEQ ID NO:1 as shown in Figure 1A and
20 1B or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO218-encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO218 gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ³²P or ³⁵S, or enzymatic labels such as alkaline
25 phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO218 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

- 30 The ESTs disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO218 sequences.

Nucleotide sequences encoding a PRO218 polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO218 polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO218 encode a protein which binds to another protein (example, where the PRO218 polypeptide functions as a receptor), the PRO218 polypeptide can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO218 polypeptide can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO218 or a receptor for PRO218. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO218 polypeptide or any of its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO218 polypeptide can be used to clone genomic DNA encoding PRO218 in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO218. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO218 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO218 introduced into the germ line of the animal

at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO218. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO218 can be used to construct a PRO218 "knock out" animal which has a defective or altered gene encoding PRO218 as a result of homologous recombination between the endogenous gene encoding PRO218 and altered genomic DNA encoding PRO218 introduced into an embryonic cell of the animal. For example, cDNA encoding PRO218 can be used to clone genomic DNA encoding PRO218 in accordance with established techniques. A portion of the genomic DNA encoding PRO218 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO218 polypeptide.

F. Anti-PRO218 Antibodies

The present invention further provides anti-PRO218 polypeptide antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

06-07-2019

15

20

25

30

0083495-04293
95240-5643009

typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against a PRO218 polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster

ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

The anti-PRO218 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the

FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

5 Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 10 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeven et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human 15 species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); 20 Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)).

25 4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a PRO218 polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

30 Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, *Nature*, 305:537-539 (1983)]. Because of the random

assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

5 Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the
10 fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

15 5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373;
20 EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No.
25 4,676,980.

 G. Uses for anti-PRO218 Antibodies

The anti-PRO218 antibodies of the present invention have various utilities. For example, anti-PRO218 antibodies may be used in diagnostic assays for PRO218 polypeptides,
30 e.g., detecting expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press,

6008345300

Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); David et al., *Biochemistry*, 13:1014 (1974); Pain et al., *J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30:407 (1982).

10 Anti-PRO218 antibodies also are useful for the affinity purification of PRO218 polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a PRO218 polypeptide are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO218 polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO218 polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO218 polypeptide from the antibody.

15 The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

25 Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

30 EXAMPLE 1: Isolation of cDNA Clones Encoding Human PRO218

A consensus DNA sequence was assembled manually relative to other EST sequences. This consensus sequence is herein designated DNA-from DNA (also referred to as DNA17411, SEQ ID NO:8) (see Figure 4). The proprietary Genentech EST sequences are

identified in Figure 4 as DNA14472 (SEQ ID NO:3) and DNA15846 (SEQ ID NO:4).

Based on the DNA-from DNA consensus sequence shown in Figure 4, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO218. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AAGTGGAGCCGGAGCCTTCC-3' (SEQ ID NO:5); and
15 reverse PCR primer 5'-TCGTTGTTTATGCAGTAGTCGG-3' (SEQ ID NO:6).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA-from dna sequence which had the following nucleotide sequence:

20 hybridization probe
5'-ATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGG-3' (SEQ ID NO:7).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO218 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB28). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al.,

Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO218 [herein designated as UNQ192 (DNA30867-1335)] (SEQ ID NO:1) and the derived protein sequence for PRO218.

The entire nucleotide sequence of UNQ192 (DNA30867-1335) is shown in Figure 1A and 1B (SEQ ID NO:1). Clone UNQ192 (DNA30867-1335) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 150-152 [Kozak et al., *supra*] and ending at the stop codon at nucleotide positions 1515-1517 (Figure 1A and 1B). The predicted polypeptide precursor is 455 amino acids long (Figure 2A). The full-length PRO218 protein shown in Figure 2A has an estimated molecular weight of about 52917 and a pI of about 9.5. Clone UNQ192 (DNA30867-1335) has been deposited with the ATCC on 28 April 1998. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO218 polypeptide suggests that PRO218 may be a novel transmembrane protein.

Still analyzing the amino acid sequence of SEQ ID NO:2, the putative signal peptide is at about amino acids 1 through 23 of SEQ ID NO:2. Transmembrane domains are potentially at about amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, and 425-444 of SEQ ID NO:2. N-glycosylation sites are at about amino acids 67, 180, and 243 of SEQ ID NO:2. Eukaryotic cobalamin-binding protein is at about amino acids 151-160 of SEQ ID NO:2. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 2: Use of PRO218-encoding DNA as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO218 as a hybridization probe.

DNA comprising the coding sequence of full-length PRO218 (as shown in Figure 1A and 1B, SEQ ID NO:1) or a fragment thereof is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO218) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO218 polypeptide-derived probe to the filters is performed in a solution of 50% formamide, 5 x SSC, 0.1% SDS,

0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2 x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1 x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO218 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO218 Polypeptides in *E. coli*

This example illustrates the preparation of unglycosylated forms of PRO218 polypeptides by recombinant expression in *E. coli*.

The DNA sequence encoding the full-length PRO218 or a fragment or variant thereof is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO218 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO218 polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the polypeptide.

EXAMPLE 4: Expression of PRO218 Polypeptides in Mammalian Cells

This example illustrates preparation of glycosylated forms of PRO218 polypeptides by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO218-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO218-encoding DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO218.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO218 DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO_4 , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 $\mu\text{Ci/ml}$ ^{35}S -cysteine and 200 $\mu\text{Ci/ml}$ ^{35}S -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO218 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO218-encoding DNA may be introduced into 293 cells transiently using the dextran sulfate method described by Sompayrac et al., *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO218 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 $\mu\text{g/ml}$ bovine insulin and 0.1 $\mu\text{g/ml}$ bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample

866240-56438003

containing expressed PRO218 polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO218 polypeptide can be expressed in CHO cells. The pRK5-PRO218 vector can be transfected into CHO cells using known reagents such as CaPO_4 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S -methionine. After determining the presence of PRO218 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO218 polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO218 polypeptide may also be expressed in host CHO cells. The PRO218-encoding DNA may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO218-encoding DNA insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO218 polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

EXAMPLE 5: Expression of a PRO218 Polypeptide in Yeast

The following method describes recombinant expression of PRO218 polypeptides in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO218 polypeptide from the ADH2/GAPDH promoter. DNA encoding the PRO218 polypeptide of interest, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO218 polypeptide. For secretion, DNA encoding the PRO218 polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO218 polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast

supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO218 polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the
5 PRO218 polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 6: Expression of PRO218 Polypeptides in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO218 polypeptides in Baculovirus-infected insect cells.

10 The PRO218-encoding DNA is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO218-encoding DNA or the desired portion of the PRO218-encoding DNA (such as the
15 sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and
20 BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 to 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford:Oxford University Press (1994).

25 Expressed poly-his tagged PRO218 polypeptide can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20
30 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with

008343-04250
055270-56438009

25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM
5 Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+} -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO218 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO218 polypeptide can
10 be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

EXAMPLE 7: Preparation of Antibodies that Bind PRO218 Polypeptides

This example illustrates the preparation of monoclonal antibodies which can specifically
15 bind to PRO218 polypeptides.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO218 polypeptide, fusion proteins containing a PRO218 polypeptide, and cells expressing recombinant PRO218 polypeptide on the cell surface. Selection of the immunogen can be
20 made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO218 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The
25 immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO218 polypeptide antibodies.

30 After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO218 polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as

dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

5

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺.

Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2×10^6 cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1×10^7 cells/ml (approx. OD₆₀₀=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 μ l) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 μ g, vol. < 10 μ l) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 μ l, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 5 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 μ l, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

10 Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 15 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 20 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red 25 starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and 30 diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified.

An aliquot of the reactions (5 μ l) were examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook

et al., *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA). Sequencing identified novel ESTs having SEQ ID NO:3 or 4.

5 Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA30867-1335		28 April 1998

10

866240 56733009
15 This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposit will be made available by ATCC
20 under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

25 The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

30 The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect

of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

866240" 56438009

WHAT IS CLAIMED IS:

1. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding a PRO218 polypeptide comprising the sequence of amino acid residues 1 or 24 through 455 of Figure 2A (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a).
- 5 2. The nucleic acid of Claim 1, wherein said DNA comprises nucleotides 150 or 219 through 1514 of the nucleotide sequence of SEQ ID NO:1.
3. The nucleic acid of Claim 1, wherein said DNA comprises the nucleotide
10 sequence of SEQ ID NO:1.
4. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit of DNA30867-1335, or (b) the complement of the DNA molecule
15 of (a).
5. The nucleic acid of Claim 4 which comprises a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in said ATCC Deposit.
- 20 6. A vector comprising the nucleic acid of any one of Claims 1 through 6.
7. The vector of Claim 6 operably linked to control sequences recognized by a host cell transformed with the vector.
- 25 8. A host cell comprising the vector of Claim 6.
9. The host cell of Claim 8, wherein said cell is a CHO cell.
10. The host cell of Claim 8, wherein said cell is an *E. coli*.
- 30 11. The host cell of Claim 8, wherein said cell is a yeast cell.
12. A process for producing a PRO218 polypeptide comprising culturing the host

cell of Claim 8 under conditions suitable for expression of said PRO218 polypeptide and recovering said PRO218 polypeptide from the cell culture.

13. Isolated native sequence PRO218 polypeptide comprising amino acid residues 1 or 24 through 455 of Figure 2 (SEQ ID NO:2).

5

14. Isolated PRO218 polypeptide encoded by the cDNA insert of the vector deposited as DNA30867-1335 with the ATCC.

15. A chimeric molecule comprising a PRO218 polypeptide fused to a heterologous amino acid sequence.

16. The chimeric molecule of Claim 15, wherein said heterologous amino acid sequence is an epitope tag sequence.

17. The chimeric molecule of Claim 15, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

18. An antibody which specifically binds to a PRO218 polypeptide.

19. The antibody of Claim 18, wherein said antibody is a monoclonal antibody.

20. An isolated nucleic acid comprising the nucleotide sequence of SEQ ID NO:3 or its complement.

21. An isolated nucleic acid comprising the nucleotide sequence of SEQ ID NO:4 or its complement.

Abstract of the Disclosure

The present invention is directed to novel polypeptides having homology to membrane regulator proteins and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous
5 polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

866240" 56488009

866240.1 SatE8009

> DNA30867 [Min]
> 0 Sites (No Sites)
> length: 2037 bp (circular)

SEQ. ID NO:1-1 CGGACGGCTG GGGGAGCGG TGGGGAGAG CGCAGTCCG GGTGACGCA CTTGGGAGAA GGCAGACGT GTGAGGGGG CTGTGGCCCC AGCGTGTGT
GCTTCGCAC CCGCTTCGC ACCCTCTC GGCCTGAGG CGAGCTGT GAGCTCTT CCGTGGCA CACTCCCG CACACCGGG TCGCACGACA

101 GACCTCGGG AGTGGAGT GAGGAGGA GCTTCTTTA CACTTCGCG TAGTTTCT CATGACTCC AGCATATCA TTACCTCCA GATACTATT
CTGAGCCCC TCACCTTCA CTTGCTCT CGGAGGAT GTGAGCGT ACTCAAGGA GTAGCTGAGG TGTATTAAT ATGAGGGT CTATGATAA
1 M S F L I D S S I H I T S Q I L P
-Start

201 TTTCGATTG GTGCGTTT CTTCATCGC CATTTCTTA AGACTATGA GATAGCTAG TATGTTGAC AGGTATCTT CTCCTGACG TTTCATTTT
AAACTTAAC CCACGAAA GAGTACCG GTTAACAAT TTCTGATAT CTATGAGT ATACACATG TCCACTAGAA GAGGCACTGC AAAGTAAAA
18 F G F G W L F F H R Q L F K D Y E I R Q Y V V Q V I P S V T F A P S

301 CTTGCACCAT GTTGGAGCT ATCATCTTTG AAATCTTAG AGTATTGAAT AGCAGCTCC GTTATTTCA CTGGAATG AACCTGTG TAATCTGCT
GACGTGGA CAACTCGAG TTAGAGAAC TTAGAACTC TCATACTTA TCGTGGAGG CAATAAAGT GACCTTTTAC TTGACACAC ATTAAGACGA
52 C T H F E L I I P E I L G V L N S S R Y F H M K N N L C V I L L

401 GATCTGTT TTCACTGTC CTTTTCAT TTGCTATT TTGAGCA ATATCGACT ACTGCATAA CAGGACTGC TTTTTCCTG TCTCTTATG
CTAGACCA AACTACCA GAAATGTA ACGATAAA TAACTCTG TATAGCTGA TGAATATT GTTCTGACG AAAAAGGAC AGAGATACC
85 I L V P H V P P Y I G Y F I V S N I R L L H K Q R L L P S C L L W

501 CTGACCTTTA TGTATTCT CTGGAACCTA GAGATCCCT TTCCATCTT CAGCCAAAA CATGGATCT TATCCATAGA ACAGCTATC AGCGGGTTG
GACTGAAAT ACATAGAA GACTTTGAT CTTCTAGGA AAGGTAGA GTGGGTTT GTACCTAGA ATAGGTATCT TGTGAGTAG TCGCCCAAC
118 L T P M Y F F W K L G D P F P I L S P K H G I L S I E Q L I S R V G

601 GTGTATGG AGTACTCT ATGCTCTTC TTCTGGAT TTGCTGTC AACCTGCCAT ACATTAAT GTCTTACTT CTCAGGAATG TACTGACAC
CACACTAAC TCACTGAG TACGAGAG AAGACTTA ACCACACAG TTGAGGGTA TGTGAATGTA CAGATGAG GAGTCTTAC ACTGACTGTG
152 V I G V T L H A L L S G F G A V N C P Y T Y M S Y P L R N V T D T

701 GATATTCTA GCCCTGGAAC GCGACTGCT GGAACATG GATATGTA TAGCAAAA GAAAGGATG GCATGCCAC GAGAGCAAT GTTCCAGAC
CCTATAGAT CCGGACTTG CCGCTAGCA CTTTGGTAC CTATCTAT ATTGTTTT CTCTCTAC CCTCTGTTA CAGGTCTTC
185 D I L A L E R R L L Q T H D M I I S K K K R N A M A R R T M F Q K

801 GGGGAGTGC ATACAAAC ATCAGTTTC TGGGAATGA TAAAGTCT TACCACTCA GCATCAGAA GTCAAAATCT TACTCTTAT CAACAGGAG
CCCCTCACG TATTGTTGG TAGTCCAAAG ACCCTTACT ATTTTACA ATGTTGAA GTAGTCTT CACTTTAGA ATGAGATAA GTTGTCTTC
218 G E V H N K P S G F W G H I K S V T T S A S G S E N L T L I Q O E V

FIGURE 1A

866240" 564E8009

901 TGGATGCTTT GGAAGATTGA AGCAGGCGC TTTTCTGGA AACAGCTGAT CTATATGCTA CCAGAGGAG AATAGANTAG TCCAAACCT TCAAGGGGAA
252 D A L E S R Q L F L E T A D L Y A A T K E R I E Y S K T P K G K
1001 ATATTATTAT TTTCTGCTGT ACTTTTCTC TATTACTGT GTTGGGAAA TTTTCATGCG TACCATCAAT ATTGTTTTG ATCGAGTTGG GAAAGCGAT
TATAAATTGA AAGAACCAG AATAAGTACA CAACCTTTT TAAAGTACCG ATGATAGCTA TAACAAAAC TAGCTCAACC CTTTTCCCTA
285 Y F N F L G Y F F S I Y C V W K I F H A T I N I V F D R V G K T D
1101 CCTGTCAAA GAGGCAITGA GATCACTGCG AATTATCTCG GAATCCAAIT TGAATGAGG TTTTGGTCCC AACCAATTTC CTTCATCTTT GTTGAATAA
GGACAGTGT CTCCTGTAAT CTAGAGACAC TTAATAGACC CTGAGTTAA ACTACACTTC AATACAGGCG TTCTGTAAG GAATGAGAA CAACCTTAT
318 P V T R G I E I T V N Y L G I Q F D V K F W S Q H I S F I L V G I I
1201 TCATGCTCAC ATCCATCAGA GGATTCCTGA TCACCTTAC CAATCTCTTT TATGCCATCT CTACCTCAAT GTCTCTCCCT GTCATTCTCC TCGTATTAGC
AGTAGCAGTG TAGGTAGTCT CCTAGGACT AGTAGAATG GTTCAGAAA ATACGGTAGA GATCTGTAAT GATGAGTTA CAGTAACAGG ACCATAATCG
352 I V T S I R G L L I T L T K F F Y A I S S K S N V I V L L L A
1301 ACAGATAATG GCGATGACT TTGTCTCTC TGTGCTGCTG ATCCGAATGA GTATGCTTTT AGAATACCC ACCATAATCA CTGAAGTCTT TGGAGAACTG
TGCTATTAC CCGTACATGA AACAGAGAG ACACGACGAC TAGCTTACT CATAGGAAA TCTTATGCGG TGGTATTAGT GACTTCAGGA ACCTCTTGAC
385 Q I M G M Y F V S V L L I R M S M P L E Y R T I I T E V L G E L
1401 CAGTCAACT TCTATCACG TTGCTTTGAT GTGATCTCC TGTGACGCG TGTCTTAGC ATACTCTCC TCTATTGGC TCACAAACAG GCACACAGAG
GTCACTTGA AGATAGTGGC AACCAACTA CACTAGAGG ACCAGTGGG AGAGAGATCG TATAGAGCCG AGATAACCG AGTCTTTCTC CTTGCTCTCT
418 Q P N P Y H R W P D V I F L V S A L S S I L F L Y L A H K Q A P E K
1501 AGCAATGCG ACCTGCTACT TAGGCTTACT ACAGCTGTT AGAGGCGAGT GCTTTCAAAA TTTAGATATA AGAGGGGGA AATATGGAAC CAGGCTTGA
TGTATTACG TGGAACTTGA ATTGGATGA TGTCTGAAA TCTCGGCTCA CAAAGTTTT AATCTATAT TCTCCCCCT TTTTACCTTG GTCCCGACT
452 Q M A P O
1601 CATTATTAA ACAACAAA TCGTATGTA GCATTTTCA CTTCTATAGC ATACTCTTC CCGTCAAGT GATPACTGA CCATGAGTAG CATCAGCCAG
GTAAATATT TGTGTTTTT ACCATACCAT CGTAAAGT GGAAGTATCG TATGAGGAG GGCAGTCCA CTATGATACT GGTACTCATC GTAGTCCGTC
1701 AACATGAG GAGAGACTTA CTCAGACAA TACTCAGAG AGAGATCCG GTGTGATAT GAGGCTGTG TAGAGCGGA GAGGAGGCAA GAACTTAAG
TTGTACTCTC CCTCTGATTT GATTTCTGTT ATGATCTGTC TGTGATGAGG CACACTATA CTCGAGCAC ATCTCGCT CTCTCGGTT CTTCGATTC
-30867.cm.p
1801 GTGAAAAATA CACTGMACT CTGGGCGAG ACATGCTAT GTTAGCTGAG CCAACACGT AGGATTTCCG TTTTAGGTT CACATGAAA AGCTTATAGC
CACTTTTTAT GTACCTTGA GACCCGTTTC TGTACAGATA CCATGACTC GCTTTGTGCA TCTTAAGCG AATATCCAA GTGTACTTT TCCAATATCG
1901 TTTGCTTGA GATTGACTCA TTAATAATCAG AGACTGAAC AAAAAAAA AGGGGGGCG CGACTCTAGA GTCGACTCG AGAAGCTTG
AAACGCACT CTAAGTGAAT AATTGATG TCTGACATG TTTTTTTT TTTTTTTTT TTTTTTTTT TTTTTTTTT TTTTTTTTT TTTTTTTTT
2001 CCGCATGCG CCACTGTT TATTGAGCT TATAATG
GGGCTACCG GGTGACAA ATACGTGA ATATTAC

FIGURE 1B

866270" 564E8009

<subunit 1 of 1, 455 aa, 1 stop
<NW: 52917, pI: 9.50, NX(S/T): 3
< 1 10 20 30 40 50 60 70
SEQ ID NO:2 MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIQRVYVQVIFSVTFPAFSCTFELIIFEILGVLNSSS
< 71 80 90 100 110 120 130 140
< RYFHWKMLCVILLILVFNVPFYIGYFIVSNIRLLHKQRLLESCLLMLTFMYFFWKLGDPPFPILSPKHGI
< 141 150 160 170 180 190 200 210
< LSIEQLISRVGIVGVTLMALLSGFGAVNCPYTYMSYFLRNVTDDILALERLLQTMDMIISKKRMAMA
< 211 220 230 240 250 260 270 280
< RRTMFQKGEVHNKPSGFWGMIKSVTTSASSENLTLIQQEVDALAEELSRQLFLETADLYATKERIEYSKT
< 281 290 300 310 320 330 340 350
< FKGKYFNELGYPFSIYCVWKIFMATINIVFDRVGKTDPTVRGIEITVNYLGIQFDVKFWSQHISFILVGI
< 351 360 370 380 390 400 410 420
< IIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMVFVSSVLLIRMSMPLEYRTIITEVLGELQFN
< 421 430 440 450
< FYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

FIGURE 2A

866240 564E8009

Signal peptide:
1-23 MSFLIDSSIMITSQILFFGFGWL

Potential transmembrane domain:
37-55 QYVQVIFSVTFAPSCTMF
81-102 VILLILVFMVPPFYIGYFIVSNI
150-168 VGIVGVTLMALLSGFGAVN
288-311 FLGYFFSIYCVWKIFMATINIVFD
338-356 FWSQHISFILVGIIIVTSTI
375-398 SSNVIVLLLAQIMGMFVSVSLLI
425-444 WFDVIFLVSAALSSILFLYLA

N-glycosylation site.
67 NSSS
180 NVTD
243 NLTL

Eukaryotic cobalamin-binding proteins
151 GVIGVTLMAL

FIGURE 2B

DNA14472:

```
>< /usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA14472 (553 bases)
< good sequence: 1-553 (553 bases)
< insert: 1-407 (407 bases), 5 regions found
<   3' pSST-AMY.1 + lk: 408-1634, 145 matches (99%), 93 consec
<   cDNA linker: 408-1634, 20 matches (100%), 20 consec
<   cDNA linker: 408-6633, 16 matches (80%), 12 consec
<   cDNA linker: 422-6633, 12 matches (80%), 7 consec, 3 gaps
<   Amylase: 462-1688, 91 matches (98%), 52 consec
CATGGGAAGTGGAGCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAG
CATCATGATTACCTCCNGANACTATTTTTGGATTGGGTGGCTTTCTTCNGCGCCAA
TGTTAAAGACTATGAGATACGTCAGTATGTGTACNGGTGATCTCTCCGTGACGTTTG
CCATTTCCTGCACCATTGTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCA
GCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCGCTGATCCTGGTTNTCA
TGGTGCCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAC
GACTGCTTTTTCCTGTCTCTTATGGCTGACCTTTATGTATTCCAG SEQ. ID NO: 3
```

DNA15846:

```
DNA15846 [ABI300.14, LIB97, YST108, Human salivary Gland] [rev comp]
good sequence: 289-920 (632 bases)
insert: 1-745 (457 bases), 5 regions found
  3' pSST-AMY.1 + lk: 746-1634, 162 matches (97%), 93 consec, 2 gaps
  cDNA linker: 746-1634, 20 matches (100%), 20 consec
  cDNA linker: 746-6633, 16 matches (80%), 12 consec
  cDNA linker: 760-6633, 12 matches (80%), 7 consec, 3 gaps
  Amylase: 800-1688, 108 matches (96%), 39 consec, 2 gaps
GTGTGGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAAATTTGGCCAAGGGTT
TCTTTNTTGAATTCGGGTTNNGNATACCTTCCCAGAAAATATTTTGGATTGGGGTA
GNTTTTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGT
GATNTTNTCCGTGACGTTTGCCATTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAAT
NTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAAT
TCTGCTGATCCTGGTTTTCATGGTGCCCTTTTACATTGGCTATTTTATTGTGAGCAATAT
CCGACTACTGCATAAACACGACTGCTTTTTCCTGTCTNTTATGGCTGACCTTTATGTA
TTTNTTNTGGAAANTAGGAGATCCCTTTCCCATTC SEQ. ID NO: 4
```

FIGURE 3

866240" 56HE8009

> DN17411 (Min)
 > 0 Sites (No Sites)
 > length: 637 bp (circular)

```

1  CATGGAGGT GGAGCGGAG CCTTCCTTAC ACTGCCATG AGTTCTCTA TGGACTECAG CATCATGAT ACCCTCCAGA TACTATTTT TGGATTTGGG
   GTACCCCTCA CTCGGCCTC GGAAGGATG TGAAGGATG TGAAGGATG GTACTACTAA TGGAGGTTCT ATGATAAAAA ACCTAAACCC
101 TGGCTTTTCT TCATGGCCA ATTGTTTAA GACTATGAGA TAGCTCTTCT CCGTGACGT TGCATTTTCT TGCACCATGT
   ACCGAAAGA AGTACCGGT TAACAAATTT CTGATACTCT ATGCAGTCAT ACAACATGTC GGCACCTGCA ACCTAAAGA ACCTGGGTAC
201 TTGAGCTCAT CATCTTTGAA ATCTTAGGAG TATTGAATAG CAGCTCCCT TATTTTCACT GGAAATGAA CCTGTGTGTA ATTCTGTGA TCCTGGTTT
   AACTCGATG GTAGAACTT TAGAATCCTC ATNACTTATC GTGAGGGCA ATAAAGTGA CCTTTACTT GGACACAT TAAGACONCT AGGACCAAAA
301 CATGGTCCCT TTTTACATG GCTATTTTAT TGTGAGCAT ATCCGACTAC TGCATAAACA ACGACTGCTT TTTTCTGTC TCTTATGGCT GACCTTTATG
   GTACCAAGCA AAAATGTAC CGTAAATA ACATCGTTA TAGCTCATG ACGTATTTGT TCTGACGAA AAGGACNG AGATACCGA CTGGAAATAC
401 TATTTCTTCT GGAATAGG AGATCCCTTT CCCATCTCA GGCAMACA TGGGATCTTA TCCATAGN ACHACTTNT NANCGGGTT GGTGTGATTC
   ATAAAGAAGA CTTTGTATCC TCTAGGAAA GGGTAAGAGT CCGGTTTTGT ACCCTAGAT AGGATATCTN TGTGCGANNA NTNGCCCCAA CCACACTAAC
501 GAGTGACTCT CATGGCTCTT CTTTCTGGAT TTGGTCTGT CAACTGCCA TACACTTACA TGTCTTACTT CCTCAGGAAT GTGACTGACA CGGATATTTCT
   CTCACCTGAG GTACCGGAA GAAAGACCTA AACCAAGCA GTTACGGGT ATGTGAATGT ACAGAATGAA GGAGTCTCTTA CACTGACTGT GCCTATAAGA
601 AGCCCTGGAA CGGCGACTGC TGCAAACCAT GGATATG -SEQ. ID NO:8
   TCGGGACCTT GCGCTGACG ACGTTTGGTA CCTATAC
  
```

FIGURE 4



587557

JCS41 U.S. PTO
08/15/98

PROVISIONAL APPLICATION COVER SHEET

(This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2))

"EXPRESS MAIL" MAILING LABEL

Number: EL029353477US Date of Deposit: May 15, 1998

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to the Box Provisional Patent Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

Typed or Printed Name: Geody Domingo

Signed: *Geody Domingo*

Docket Number:	P-66136/WHD/DAV (G-tech Docket: PR1381)	Type a plus sign (+) inside this box -	+
----------------	--	--	---

INVENTOR(s)/APPLICANT(s)					
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)		
WOOD	WILLIAM	I.	San Mateo, California		
TITLE OF THE INVENTION (280 characters max)					
NOVEL POLYPEPTIDES HAVING SEQUENCE IDENTITY WITH IL-17 AND NUCLEIC ACIDS ENCODING THE SAME					
CORRESPONDENCE ADDRESS					
Walter H. Dreger, Registration No. 24, 190 FLEHR HOHBACH TEST ALBRITTON & HERBERT LLP, 4 Embarcadero Center, Suite 3400, San Francisco					
STATE	CA	ZIP CODE	94111	COUNTRY	US
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification	Number of Pages	39	<input type="checkbox"/>	Small Entity Statement
<input checked="" type="checkbox"/>	Drawings	Number of Sheets	5	<input type="checkbox"/>	Other (specify)
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the Provisional filing fees			PROVISIONAL FILING FEE AMOUNT (\$)	\$150
<input checked="" type="checkbox"/>	The Commissioner is hereby authorized to charge any underpayment or overpayment to Deposit Account Number.			06-1300 (Order No. P-66136/WHD/DAV)	

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒

No

☐

Yes, the name of the U.S. Government Agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE: *Walter H. Dreger*

Date

15 May 1998

TYPED or PRINTED NAME Walter H. DregerREGISTRATION NO
(if appropriate)

24,190

☐

Additional inventors are being named on separately numbered sheets attached hereto.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden Hour Statement: This form is estimated to take .2 hours to complete. Times will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Assistant Commissioner for Patents, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.



PATENT
Attorney Docket No.: P-66136/WHD/DAV
(Genentech Docket No.:PR1381)

NOVEL POLYPEPTIDES HAVING SEQUENCE IDENTITY WITH IL-17 AND
NUCLEIC ACIDS ENCODING THE SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel
5 DNA and to the recombinant production of novel polypeptides having sequence identity with
the cytokine IL-17, designated herein as "PRO1031" polypeptides.

BACKGROUND OF THE INVENTION

It has been reported that the cytokine interleukin 17 (IL-17) stimulates epithelial,
10 endothelial, and fibroblastic cells to secrete cytokines such as IL-6, IL-8, and granulocyte-
colony-stimulating factor, as well as prostaglandin E2. Moreover, it has been shown that
when cultured in the presence of IL-17, fibroblasts could sustain proliferation of CD34+
preferential maturation into neutrophils. Thus it has been suggested that IL-17 constitutes
an early initiator of the T cell-dependent inflammatory reaction and/or an element of the
15 cytokine network that bridges the immune system to hematopoiesis. See, Yao, et al., *J.*
Immunol., 155(12):5483-5486 (1995); Fossiez, et al., *J. Exp. Med.*, 183(6):2593-2603
(1996); Kennedy, et al., *J. Interferon Cytokine Res.*, 16(8):611-617 (1996). Thus, proteins
related to IL-17 are of interest.

More generally, all novel proteins are of interest. Extracellular proteins play an
20 important role in the formation, differentiation and maintenance of multicellular organisms.
The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction
with other cells, is typically governed by information received from other cells and/or the
immediate environment. This information is often transmitted by secreted polypeptides (for
instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors,
25 neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell
receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules
normally pass through the cellular secretory pathway to reach their site of action in the
extracellular environment.

Secreted proteins have various industrial applications, including pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents.

5 Efforts are being undertaken by both industry and academia to identify new, native secreted proteins, particularly those related to IL-17. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. 10 Patent No. 5,536,637)]. The results of such efforts are presented herein.

SUMMARY OF THE INVENTION

Applicants have identified a cDNA clone that encodes a novel polypeptide having sequence identity with IL-17, wherein the polypeptide is designated in the present application 15 as "PRO1031".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO1031 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO1031 polypeptide having amino acid residues 1 through 180 of Figure 2 (SEQ ID NO:2), or is complementary to such encoding nucleic acid 20 sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. The isolated nucleic acid sequence may comprise the cDNA insert of the vector deposited on 14 May 1998 with the ATCC as DNA59294-1381 which includes the nucleotide sequence encoding PRO1031.

In another embodiment, the invention provides a vector comprising DNA encoding 25 a PRO1031 polypeptide. A host cell comprising such a vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing PRO1031 polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of PRO1031 and recovering PRO1031 from the cell culture.

In another embodiment, the invention provides isolated PRO1031 polypeptide. In 30 particular, the invention provides isolated native sequence PRO1031 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 180 of Figure 2 (SEQ ID NO:2). Optionally, the PRO1031 polypeptide is obtained or is obtainable

by expressing the polypeptide encoded by the cDNA insert of the vector deposited on 14 May⁵ 1998 with the ATCC as DNA59294-1381.

In another embodiment, the invention provides chimeric molecules comprising a PRO1031 polypeptide fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises a PRO1031 polypeptide fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which specifically binds to a PRO1031 polypeptide. Optionally, the antibody is a monoclonal antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

10 Figure 1 shows a nucleotide sequence (SEQ ID NO:1) containing the nucleotide sequence of a native sequence PRO1031 cDNA (nucleotides 42-581 of SEQ ID NO:1), wherein the nucleotide sequence (SEQ ID NO:1) is a clone designated herein as "UNQ516" and/or "DNA59294-1381". Also presented is the position of the initiator methionine residue, circled, and the stop codon, circled. The complementary strand and deduced amino acid
15 sequence are also shown.

20 Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from nucleotides 42-581 of SEQ ID NO:1. Also shown in Figure 2 are the signal peptide, an N-glycosylation site, a region having sequence identity with IL-17, the molecular weight, and approximate pI.

25 Figures 3A-3C show an alignment of nucleotide sequences from a variety of expressed sequence tags as well as a consensus nucleotide sequence derived therefrom designated "DNA-from dna" shown as "<consen01>".

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

25 I. Definitions

The terms "PRO1031 polypeptide" and "PRO1031" when used herein encompass native sequence PRO1031 and PRO1031 polypeptide variants (which are further defined herein). The PRO1031 polypeptides may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

30 A "native sequence PRO1031 polypeptide" comprises a polypeptide having the same amino acid sequence as a PRO1031 polypeptide derived from nature. Such native sequence PRO1031 polypeptide can be isolated from nature or can be produced by recombinant or

synthetic means. The term "native sequence PRO1031 polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of a PRO1031 polypeptide (e.g., soluble forms containing for instance, an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of a PRO1031 polypeptide. In one embodiment of the invention, the native sequence PRO1031 polypeptide is a full-length or mature native sequence PRO1031 polypeptide comprising amino acids 1 or 21 through 180 of Figure 2 (SEQ ID NO:2). Optionally, the PRO1031 polypeptide is obtained or obtainable by expressing the polypeptide encoded by the cDNA insert of the vector DNA59294-1381 deposited on 14 May 1998 with the ATCC.

It will be understood that any transmembrane domain identified for the PRO1031 polypeptide of the present invention is identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified.

"PRO1031 variant" means an active PRO1031 polypeptide as defined below having at least about 80% amino acid sequence identity with the PRO1031 polypeptide having the deduced amino acid sequence shown in Figure 2 (SEQ ID NO:2) for a full-length or mature native sequence PRO1031 polypeptide. Such PRO1031 polypeptide variants include, for instance, PRO1031 polypeptides wherein one or more amino acid residues are added, substituted or deleted, at the N- or C-terminus or within the sequence of Figure 2 (SEQ ID NO:2). Ordinarily, a PRO1031 polypeptide variant will have at least about 80% amino acid sequence identity, preferably at least about 85% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity and even more preferably at least about 95% amino acid sequence identity with the amino acid sequence of Figure 2 (SEQ ID NO:2), with or without the signal peptide. The variants provided herein exclude the polypeptides and nucleic acids described herein which the novel polypeptides have identity with and which are already known in the art.

"Percent (%) amino acid sequence identity" with respect to the PRO1031 amino acid sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a PRO1031 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence

identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Similarity can be determined using the same programs.

5 "Percent (%) nucleic acid sequence identity" with respect to the PRO1031 sequence identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO1031 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved
10 in various ways that are within the skill in the art, for instance, using publicly available computer software such as ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Similarity can be determined using the same programs.

15 Most of the WU-BLAST-2 search parameters as used herein were set to the default values. The adjustable parameters were set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. The HSP S and HSP S2 parameters, which are dynamic values used by BLAST-2, are established by the program itself depending upon the composition of the sequence of interest
20 and composition of the database against which the sequence is being searched. However, the values may be adjusted to increase sensitivity. A % sequence identity value is determined by the fraction of matching identical residues divided by the total number of residues of the "longer" sequence. The "longer" sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-BLAST-2 to maximize the alignment score are
25 ignored).

The term "epitope tagged" where used herein refers to a chimeric polypeptide comprising a PRO1031 polypeptide, or domain sequence thereof, fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody may be made, or which can be identified by some other agent, yet is short
30 enough such that it does not interfere with the activity of the PRO1031 polypeptide. The tag polypeptide preferably is also fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid

residues and usually between about 8 to about 50 amino acid residues (preferably, between about 10 to about 20 residues).

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO1031 polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO1031 polypeptide-encoding nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO1031 polypeptide-encoding nucleic acid. An isolated PRO1031 polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated PRO1031 polypeptide-encoding nucleic acid molecules therefore are distinguished from the PRO1031 polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated PRO1031 polypeptide-encoding nucleic acid molecule includes PRO1031 polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express PRO1031 polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates

in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not
5 have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers single anti-PRO1031 polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing
10 antibodies) and anti-PRO1031 antibody compositions with polyepitopic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

15 "Active" or "activity" for the purposes herein refers to form(s) of PRO1031 which retain the biologic and/or immunologic activities of native or naturally-occurring PRO1031 polypeptide.

The terms "treating", "treatment" and "therapy" as used herein refer to curative therapy, prophylactic therapy, and preventative therapy.

20 The term "mammal" as used herein refers to any mammal classified as a mammal, including humans, cows, horses, dogs and cats. In a preferred embodiment of the invention, the mammal is a human.

II. Compositions and Methods of the Invention

25 A. Full-length PRO1031 Polypeptide

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1031. In particular, Applicants have identified and isolated cDNA encoding a PRO1031 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment
30 computer programs, Applicants found that various portions of the PRO1031 polypeptide have sequence identity with IL-17 and CTLA-8. Accordingly, it is presently believed that

PRO1031 polypeptide disclosed in the present application is a newly identified member of the cytokine family and thus may be involved in inflammation and/or the immune system.

B. PRO1031 Variants

5 In addition to the full-length native sequence PRO1031 polypeptide described herein, it is contemplated that PRO1031 variants can be prepared. PRO1031 variants can be prepared by introducing appropriate nucleotide changes into the PRO1031-encoding DNA, or by synthesis of the desired PRO1031 polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO1031 polypeptide,
10 such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO1031 or in various domains of the PRO1031 polypeptide described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for
15 instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO1031 polypeptide that results in a change in the amino acid sequence of the PRO1031 polypeptide as compared with the native sequence PRO1031. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO1031 polypeptide.
20 Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO1031 polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having
25 similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in any of the *in vitro* assays described in the
30 Examples below.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed

mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO1031-encoding variant DNA.

5 Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation
10 of the variant. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

15 C. Modifications of PRO1031

Covalent modifications of PRO1031 polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO1031 polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of a PRO1031 polypeptide.
20 Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO1031 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO1031 antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such
25 as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the
30 α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86

(1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO1031 polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO1031 polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO1031 polypeptide.

Addition of glycosylation sites to PRO1031 polypeptides may be accomplished by altering the amino acid sequence thereof. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO1031 polypeptide (for O-linked glycosylation sites). The PRO1031 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO1031 polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO1031 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO1031 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO1031 comprises linking the PRO1031 polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol, polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

PRO1031 polypeptides of the present invention may also be modified in a way to form chimeric molecules comprising a PRO1031 polypeptide fused to another, heterologous

polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of a PRO1031 polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO1031 polypeptide. The presence of such epitope-tagged forms of a PRO1031 polypeptide can be detected using an antibody
5 against the tag polypeptide. Also, provision of the epitope tag enables the PRO1031 polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of a PRO1031 polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the
10 chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan
15 et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and
20 the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

The PRO1031 polypeptide of the present invention may also be modified in a way to form a chimeric molecule comprising a PRO1031 polypeptide fused to a leucine zipper. Various leucine zipper polypeptides have been described in the art. See, e.g., Landschulz
25 et al., Science 240:1759 (1988); WO 94/10308; Hoppe et al., FEBS Letters 344:1991 (1994); Maniatis et al., Nature 341:24 (1989). It is believed that use of a leucine zipper fused to a PRO1031 polypeptide may be desirable to assist in dimerizing or trimerizing soluble PRO1031 polypeptide in solution. Those skilled in the art will appreciate that the leucine zipper may be fused at either the N- or C-terminal end of the PRO1031 molecule.

30

D. Preparation of PRO1031

The description below relates primarily to production of PRO1031 by culturing cells transformed or transfected with a vector containing PRO1031 polypeptide encoding nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO1031 polypeptides. For instance, the PRO1031 sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of PRO1031 polypeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce a full-length PRO1031 polypeptide.

1. Isolation of DNA Encoding PRO1031

DNA encoding a PRO1031 polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the PRO1031 mRNA and to express it at a detectable level. Accordingly, human PRO1031-encoding DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO1031-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to a PRO1031 polypeptide or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO1031 is to use PCR methodology [Sambrook et al., *supra*; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened.

Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P -labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Included in this invention are nucleic acids having at least about 70% to about 75% nucleic acid sequence identity with the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1), with or without the untranslated portions and with or without the region encoding the signal peptide. Ordinarily, in another embodiment, the nucleic acids will have at least about 80% nucleic acid sequence identity, preferably at least about 85% nucleic acid sequence identity, more preferably at least about 90% nucleic acid sequence identity and even more preferably at least about 95% nucleic acid sequence identity with the nucleic acid sequence of SEQ ID NO:1, with or without the untranslated portions and with or without the region encoding the signal peptide.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined through sequence alignment using computer software programs such as ALIGN, DNASTar, and INHERIT which employ various algorithms to measure homology.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO1031 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO₄ and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., *supra*, or electroporation is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. Infection with 5 *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., *Gene*, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, *Virology*, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 10 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., *J. Bact.*, 130:946 (1977) and Hsiao et al., *Proc. Natl. Acad. Sci. (USA)*, 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for 15 transforming mammalian cells, see Keown et al., *Methods in Enzymology*, 185:527-537 (1990) and Mansour et al., *Nature*, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, 20 Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635).

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO1031-encoding vectors. *Saccharomyces* 25 *cerevisiae* is a commonly used lower eukaryotic host microorganism.

Suitable host cells for the expression of glycosylated PRO1031 are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples 30 include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and

Chasin, Proc. Natl. Acad. Sci. USA, **77**:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., **23**:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding the desired PRO1031 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The desired PRO1031 polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO1031-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for

a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO1031-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO1031-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the PRO1031 polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase,

3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO1031 transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding a PRO1031 polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO1031 coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO1031.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO1031 polypeptides in recombinant vertebrate cell culture are described in Gething et al.,

Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO1031 polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO1031-encoding DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO1031 may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO1031 polypeptides can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO1031 from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein

A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO1031 polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification
5 step(s) selected will depend, for example, on the nature of the production process used and the particular PRO1031 polypeptide produced.

E. Uses for PRO1031

10 Nucleotide sequences (or their complement) encoding PRO1031 polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO1031-encoding nucleic acid will also be useful for the preparation of PRO1031 polypeptides by the recombinant techniques described herein.

15 The full-length DNA59294-1381 nucleotide sequence (SEQ ID NO:1) or the full-length native sequence PRO1031 nucleotide sequence, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO1031 gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of PRO1031 or PRO1031 from other species) which have a desired sequence identity to the PRO1031 nucleotide sequence disclosed in Figure 1 (SEQ ID NO:1). Optionally, the length of the
20 probes will be about 20 to about 50 bases. The hybridization probes may be derived from the UNQ516 (DNA59294-1381) nucleotide sequence of SEQ ID NO:1 as shown in Figure 1 or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO1031-encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO1031 gene using the known DNA sequence to
25 synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ³²P or ³⁵S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO1031 gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which
30 members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO1031 sequences.

Nucleotide sequences encoding a PRO1031 polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO1031 polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided
5 herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO1031 encode a protein which binds to another protein (example, where the PRO1031 polypeptide functions as a receptor), the PRO1031
10 polypeptide can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO1031 polypeptide can be used to isolate correlative ligand(s). Screening assays
15 can be designed to find lead compounds that mimic the biological activity of a native PRO1031 or a receptor for PRO1031. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats,
20 including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO1031 polypeptide or any of its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic
25 animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO1031 polypeptide can be used to clone genomic DNA encoding PRO1031 in accordance with
30 established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO1031. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and

are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO1031 transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO1031 introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO1031. Such animals can be used as
5 tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

10 Alternatively, non-human homologues of PRO1031 can be used to construct a PRO1031 "knock out" animal which has a defective or altered gene encoding PRO1031 as a result of homologous recombination between the endogenous gene encoding PRO1031 and altered genomic DNA encoding PRO1031 introduced into an embryonic cell of the animal. For example, cDNA encoding PRO1031 can be used to clone genomic DNA encoding
15 PRO1031 in accordance with established techniques. A portion of the genomic DNA encoding PRO1031 can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination
20 vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*,
25 E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals
30 can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO1031 polypeptide.

PRO1031 polypeptides of the present invention which possess biological activity⁵ related to that of IL-17 may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO1031 polypeptides of the present invention for such purposes.

PRO1031 can be used in assays with the polypeptides to which they have identity with
5 to determine the relative activities. The results can be applied accordingly.

F. Anti-PRO1031 Antibodies

The present invention further provides anti-PRO1031 polypeptide antibodies.
Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and
10 heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO1031 antibodies of the present invention may comprise polyclonal
antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan.
15 Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of
an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or
adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal
injections. The immunizing agent may include the PRO1031 polypeptide or a fusion protein
thereof. It may be useful to conjugate the immunizing agent to a protein known to be
20 immunogenic in the mammal being immunized. Examples of such immunogenic proteins
include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine
thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed
include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A,
synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one
25 skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO1031 antibodies may, alternatively, be monoclonal antibodies.
Monoclonal antibodies may be prepared using hybridoma methods, such as those described
30 by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster,
or other appropriate host animal, is typically immunized with an immunizing agent to elicit

lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO1031 polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Rockville, Maryland. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against a PRO1031 polypeptide. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, *supra*]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

5 The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

10 The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression
15 vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No.
20 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

25 The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue
30 or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Humanized Antibodies

5 The anti-PRO1031 antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include
10 human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies
15 may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus
20 sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally,
25 a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the
30 corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact

human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)).

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for a PRO1031 polypeptide, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-

transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention.

- 5 Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents.
- 10 For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

15 G. Uses for anti-PRO1031 Antibodies

- The anti-PRO1031 antibodies of the present invention have various utilities. For example, anti-PRO1031 antibodies may be used in diagnostic assays for PRO1031 polypeptides, *e.g.*, detecting expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding
- 20 assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable
- 25 moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962);
- 30 David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO1031 antibodies also are useful for the affinity purification of PRO1031 polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a PRO1031 polypeptide are immobilized on a suitable support, such as a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO1031 polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO1031 polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO1031 polypeptide from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Rockville, Maryland.

EXAMPLE 1: Isolation of cDNA Clones Encoding Human PRO1031

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank, Merck/Wash U.) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>).

An initial consensus DNA sequence was assembled relative to other EST sequences using phrap. The initial consensus DNA sequence was extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. Figure 3 shows the assembly, wherein the initial consensus sequence is shown as "DNA47332.init" and the extended consensus sequence is shown as
5 "<consen01>" (also sometimes referred to as DNA47332).

Based on the consensus sequences assembled herein and other information provided herein, Merck W74558 (clone 344649) was further examined. DNA sequencing gave the full-length DNA sequence for PRO1031 [herein designated as UNQ516 (DNA59294-1381)] (SEQ ID NO:1) and the derived protein sequence for PRO1031.

10 The entire nucleotide sequence of UNQ516 (DNA59294-1381) is shown in Figure 1 (SEQ ID NO:1). Clone UNQ516 (DNA59294-1381) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 [Kozak et al., *supra*] and ending at the stop codon at nucleotide positions 582-584 (Figure 1). The predicted polypeptide precursor is 180 amino acids long (Figure 2). The full-length PRO1031 protein
15 shown in Figure 2 has an estimated molecular weight of about 20437 and a pI of about 9.58. Clone UNQ516 (DNA59294-1381) has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1031 polypeptide suggests
20 that it is a novel cytokine.

Still analyzing the amino acid sequence of SEQ ID NO:2, the putative signal peptide is at about amino acids 1-20 of SEQ ID NO:2. An N-glycosylation site is at about amino acids 75-78 of SEQ ID NO:2. A region having sequence identity with IL-17 is at about amino acids 96-180. The corresponding nucleotides can be routinely determined given the
25 sequences provided herein.

EXAMPLE 2: Use of PRO1031-encoding DNA as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO1031 as a hybridization probe.

30 DNA comprising the coding sequence of full-length PRO1031 (as shown in Figure 1, SEQ ID NO:1) or a fragment thereof is employed as a probe to screen for homologous

DNAs (such as those encoding naturally-occurring variants of PRO1031) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO1031 polypeptide-derived probe to the filters is performed in a solution of 50% formamide, 5 x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2 x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1 x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO1031 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO1031 Polypeptides in *E. coli*

This example illustrates the preparation of unglycosylated forms of PRO1031 polypeptides by recombinant expression in *E. coli*.

The DNA sequence encoding the full-length PRO1031 or a fragment or variant thereof is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO1031 coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate

a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO1031 polypeptide can then be purified using
5 a metal chelating column under conditions that allow tight binding of the polypeptide.

EXAMPLE 4: Expression of PRO1031 Polypeptides in Mammalian Cells

This example illustrates preparation of glycosylated forms of PRO1031 polypeptides by recombinant expression in mammalian cells.

10 The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO1031-encoding DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO1031-encoding DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO1031.

15 In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO1031 DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmapaya et al., Cell, 31:543 (1982)] and dissolved in 500 μ l of 1 mM
20 Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO_4 , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with
25 serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ^{35}S -cysteine and 200 μ Ci/ml ^{35}S -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel
30 may be dried and exposed to film for a selected period of time to reveal the presence of PRO1031 polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO1031-encoding DNA may be introduced into 293 cells transiently using the dextran sulfate method described by Sompanyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO1031 DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO1031 polypeptide can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO1031 polypeptide can be expressed in CHO cells. The pRK5-PRO1031 vector can be transfected into CHO cells using known reagents such as CaPO_4 or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S -methionine. After determining the presence of PRO1031 polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO1031 polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO1031 polypeptide may also be expressed in host CHO cells. The PRO1031-encoding DNA may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO1031-encoding DNA insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO1031 polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+} -chelate affinity chromatography.

30 EXAMPLE 5: Expression of a PRO1031 Polypeptide in Yeast

The following method describes recombinant expression of PRO1031 polypeptides in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO1031 polypeptide from the ADH2/GAPDH promoter. DNA encoding the PRO1031 polypeptide of interest, a selected signal peptide and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of the PRO1031 polypeptide. For secretion, DNA encoding the PRO1031 polypeptide can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, the yeast alpha-factor secretory signal/leader sequence, and linker sequences (if needed) for expression of the PRO1031 polypeptide.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO1031 polypeptide can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing the PRO1031 polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 6: Expression of PRO1031 Polypeptides in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO1031 polypeptides in Baculovirus-infected insect cells.

The PRO1031-encoding DNA is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the PRO1031-encoding DNA or the desired portion of the PRO1031-encoding DNA (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 to 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications.

Viral infection and protein expression is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford:Oxford University Press (1994).

Expressed poly-his tagged PRO1031 polypeptide can then be purified, for example, by Ni^{2+} -chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl_2 ; 0.1 mM EDTA; 10% Glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% Glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+} -NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% Glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or western blot with Ni^{2+} -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO1031 polypeptide are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO1031 polypeptide can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

25 EXAMPLE 7: Preparation of Antibodies that Bind PRO1031 Polypeptides

This example illustrates the preparation of monoclonal antibodies which can specifically bind to PRO1031 polypeptides.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO1031 polypeptide, fusion proteins containing a PRO1031 polypeptide, and cells expressing recombinant PRO1031 polypeptide on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO1031 immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO1031 polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO1031 polypeptide. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO1031 polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against a PRO1031 polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO1031 polypeptide monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Material

ATCC Dep. No.

Deposit Date

DNA59294-1381

14 May 1998

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposit will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding a PRO1031 polypeptide comprising the sequence of amino acid residues 1 or 21 through 180 of Figure 2 (SEQ ID NO:2), or (b) the complement of the DNA molecule of (a).
- 5 2. The nucleic acid of Claim 1, wherein said DNA comprises nucleotides 42 or 102 through 581 of SEQ ID NO:1.
3. The nucleic acid of Claim 1, wherein said DNA comprises the nucleotide
10 sequence of SEQ ID NO:1.
4. Isolated nucleic acid comprising DNA having at least an 80% sequence identity to (a) a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in ATCC Deposit of DNA59294-1381, or (b) the complement of the DNA molecule
15 of (a).
5. The nucleic acid of Claim 4 which comprises a DNA molecule encoding the same mature polypeptide encoded by the human protein cDNA in said ATCC Deposit.
- 20 6. A vector comprising the nucleic acid of any one of Claims 1 through 5.
7. The vector of Claim 6 operably linked to control sequences recognized by a host cell transformed with the vector.
- 25 8. A host cell comprising the vector of Claim 7.
9. The host cell of Claim 8, wherein said cell is a CHO cell.
10. The host cell of Claim 8, wherein said cell is an *E. coli*.
- 30 11. The host cell of Claim 8, wherein said cell is a yeast cell.

12. A process for producing a PRO1031 polypeptide comprising culturing the host cell of Claim 8 under conditions suitable for expression of said PRO1031 polypeptide and recovering said PRO1031 polypeptide from the cell culture.

13. An isolated polypeptide comprising amino acid residues 1 or 21 through 180
5 of Figure 2 (SEQ ID NO:2).

14. Isolated PRO1031 polypeptide encoded by the cDNA insert of the vector deposited with the ATCC as DNA59294-1381.

10 15. A chimeric molecule comprising a PRO1031 polypeptide fused to a heterologous amino acid sequence.

16. The chimeric molecule of Claim 15, wherein said heterologous amino acid sequence is an epitope tag sequence.

15 17. The chimeric molecule of Claim 15, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

18. An antibody which specifically binds to a PRO1031 polypeptide.

20 19. The antibody of Claim 18, wherein said antibody is a monoclonal antibody.

Abstract of the Disclosure

The present invention is directed to novel polypeptides having sequence identity with IL-17 and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous
5 polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500

865750-64558005

SEQ. ID NO:1

1 AGCGGGGAG CAGTGCAGG CTGACCTTC AGCTTGGGG ATGACTGAG TGGTCTCT TCTTACCTT TCGATCTCC TGGGCTGGG
TCCGCCGCTC GTCGACCTC GACTGGAAG TCGAACGCC TTACTGACC GAGTGTGG ACACACAGA AGATGGTAA AGGTAGAGG ACCCGACCC
M D W P H N L L F L L T I S I F L C L C

101 CCAGCCAGG AGCCCAAAA GCNAGAGAA GGGCAAGG CGCCTGGG CCTGGCCCC TGGCCCTCAC CAGTGCAC TCGACCTGT GTACGGATG
GTCGGGTCC TCGGGTTTT GTTCTCTT CCGCTTCC CGCGACCG GGGACGGG GTCACGGTG ACCTGGACCA CAGTGCCTAC
21 Q P R S P K S K R K G Q G R P G P L A P G P H Q V P L D L V S R M

201 AAACCGATG CCGCATGGA GAGATATAG AGGAACATG AGGATATGT GCGCCAGCTG AGGACAGCT CAGAGCTGG CCAGCAAG TGTAGGTCA
TTTGGCATAC GCGGTACTC CCTCATACT TCTTTAGC TCTCTACCA CCGGTGAC TCTTTGTA GTCTGACC GGTCTCTTC ACATCCAGT
34 K P Y A R M E S Y E R N I E M V A Q L R N S S E L A Q R K C E V N

301 ACTTGCAGT GTGATGTCC AACAGAGA GCTGTCTCC CTGGGCTAC AGCATCAAC AGACCCGAG CGTATCCC GTGACCTGC CGAGGGCAG
TGACGTGA CACTACAGG TTGTTCTCT CGGACAGG GACCCCGATG TGTAGTTG TGTGGGTG GGCATGGG CACCTGGAG GCTTCCGTG
88 L Q L W H S N K R S L S P W G Y S I N H D P S R I P V D L P E A R

401 GTGCTGTGT CTGGGTGTG TGAACCTT CACATGAG GAGGACCA GCATGTGAG CTGCGGTG TTCCTGTGG CCGCCGCTC
CAGCGACCA GACCGACAC ACTTGGGA GTGTAGCT CTCTGGGT CGTACACT GCACGACAC AAGTGGTCC AAGGACAC GCGCGCGAG
121 C L C L G C V N P F T H Q E D R S M V S V P V F S Q V P V R R L

501 TCCCGCCAC CCGCTCCAC AGGCTTTC CCGCAGCGG CAGTATGA GACCATGCT GTGGGTGCA CTTCATCTT GTATCACC TGGCCAGAA
ACGGGCGGTG GCGGGCGGTG TCCGGGAAG GCGGTCCGCG GTGAGTACT CTGTAGCTT CAGTACGT GACATGAGAA GACTTATG ACCTGTCTT
154 C P P P R T G P C R Q R A V M E T I A V G C T C I F O

601 GCCAGCCAG CAGCCGAGA CCATCTCTT TGCACCTTG TCCCAAGAA GGCCTATGA AAGTAAAC TCACTTTGA AAGCAG
CGCTCCGCTC GTCGGCTCT GCTAGGAGA ACCTGGAAAC ACCTTCTTT CCGATACTT TCACTTTG ACTGAAACT TTCCTTC

FIGURE 1

205T50-64558885

SEQ. ID NO.2

<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSRKKGQGRPGPLAPGPHQVPLDLYSRMKPYARME
EYERNIEEMVAQLRNSSSELAQRKCEVNLQLMMSKRSLSPMGYSINHDPSRIIPVDLPEAR
CLCLGCVNPFMTQEDRSMSVVPVFSQVPTRRRLCPPPPRTGFCRQRAVMTETINVGCTCIF

signal peptide:
MDWPHNLLFLLTISIFLGLG
1-20

N-glycosylation site.

75 NSSE to IL-17

Homologous region
96-180.

FIGURE 2


```

W74558      204 GCGCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGGCCAGCT
AA033733    173 GCGCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATGGTGGCCAGCT
DNA47332.init 1      GGAACATCGAGGAGATGTTGGCCAGCT
W74664.RC   1      GGAACATCGAGGAGATGTTGGCCAGCT
               ++++++
<consen01>  348 GCGCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATGTTGGCCAGCT

AA476704    282 GAGGAACAGCTCAGA-GCTGGCCAGAGAAA-GTGTGAGGTCAA-CTTGC
2967181     279 GAGGAACAGCTCAGA-GCTGGCCAGAGAAA-GTGTGAGGT
W74558      254 GAGGAACAGCTCANAAGCTGGCCAGAGAAA-GTGTGAGGTCAA-CTTGC
AA033733    223 GAGGAACAGCTCAGA-GCTGGCCAGAGAAAATGTGTGAGGTCAA-CTTGC
DNA47332.init 30  GAGGAACAGTTCAGA-GCTGGCCAGAGAAA-GTGTGAGGTCAA-CTTGC
W74664.RC   30  GAGGAACAGTTCAGA-GCTGGCCAGAGAAA-GTGTGAGGTCAA-CTTGC
AA443286.RC 1      GAGGTCAA-CTTGC
               ++++++
<consen01>  398 GAGGAACAGTTCAGA GCTGGCCAGAGAAA GTGTGAGGTCAA CTTGC

AA476704    329 AGCTGTGGATGTCCAA-CAA-GA-GGAGCCTGTCTCCCT-GGGGCTACA-
W74558      302 AGCTGTGGATGTCCAA-CAA-GAAGGAGCCTGTCTCCCTTGGGGCTACAA
AA033733    271 AGCTGTGGATGTCCAA-CAA-GA-GGAGCCTGTCTCCCT-GGGGCTACA-
DNA47332.init 78  AGCTGTGGATGTCCAAACAAAGA-GGAGCCTGTCTCCCT-GGGGCTACA-
W74664.RC   78  AGCTGTGGATGTCCAAACAAAGA-GGAGCCTGTCTCCCT-GGGGCTACA-
AA443286.RC 14  AGCTGTGGATGTCCAA-CAA-GA-GGAGCCTGTCTCCCT-GGGGCTACA-
               ++++++
<consen01>  445 AGCTGTGGATGTCCAA CAA GA GGAGCCTGTCTCCCT GGGGCTACA

AA476704    374 GCAT
W74558      350 GCATCAACCACCGACCCC-AGCCGTATCCCCGTGGGACC-TTGCCGGGAC
AA033733    316 GCATCAACCAC-GACCCC-AGCCGTATCCCCGTGGACCTGCCGGAGNG--
DNA47332.init 125 GCATCAACCAC-GACCCC-AGCCGTATCCCCGTGGACCT-CCGGAGGCAC
W74664.RC   125 GCATCAACCAC-GACCCCNAGCCGTATCCCCGTGGACCT-CCGGAGGCAC
AA443286.RC 59  GCATCAACCAC-GACCCC-AGCCGTATCCCCGTGGACCT-CCGGAGGCAC
               ++++++
<consen01>  490 GCATCAACCAC GACCCC AGCCGTATCCCCGTGGACCT CCGGAGGCAC

W74558      398
AA033733    362 GTNGCCTGTGTCTGGGCT-GTGTGAACCCCTTCACCATGCAGGAGGACCG
DNA47332.init 172 GGTGCCTGTGTCTGGGCTTGTGTGAACCCCTTCACCATGCAGGAGGACCG
W74664.RC   173 GGTGCCTGTGTCTGGGCTTGTGTGAACCCCTTCACCATGCAGGAGGACCG
AA443286.RC 106 GGTGCCTGTGTCTGGGCT-GTGTGAACCCCTTCACCATGCAGGAGGACCG
               ++++++
<consen01>  537 GGTGCCTGTGTCTGGGCTTGTGTGAACCCCTTCACCATGCAGGAGGACCG

AA033733    411 CAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGGTTCCTG
DNA47332.init 222 CAGCATGGTGAGCGTGCCGGTGTTCAGCCAGG-TTCCTGTGCGCCGCGCGC
W74664.RC   223 CAGCATGGTGAGCGTGCCGGTGTTCAGCCAGG-TTCCTGTGCGCCGCGCGC
AA443286.RC 155 CAGCATGGTGAGCGTGCCGGTGTTCAGCCAGG-TTCCTGTGCGCCGCGCGC
               ++++++
<consen01>  587 CAGCATGGTGAGCGTGCCGGTGTTCAGCCAGG TTCCTGTGCGCCGCGCGC

DNA47332.init 271 CTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCAT
W74664.RC   272 CTCTGCCCCGCCACCGCCCCGCACAGG-CCTTGCC-CCAGCGCGCAGTCAT
AA443286.RC 204 CTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCAT
               ++++++

```

FIGURE 3B

```

<consen01>      636 CTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCAT
DNA47332.init    321 GGAGACCATCGCTGTGGGCTGCACCTGCATCTTCTGAATCGACCTGGCCC
W74664.RC        320 GGAGACCATCGCTGTGGGCTGCACCTGCATCTTCTGAATCGACCTGGCCC
AA443286.RC      254 GGAGACCATCGCTGTGGGCTGCACCTGCATCTTCTGAATC-ACCTGGCCC
                  ++++++
<consen01>      686 GGAGACCATCGCTGTGGGCTGCACCTGCATCTTCTGAATCGACCTGGCCC
DNA47332.init    371 AGAAGCCAGGCCAGCAGCCCCGAGACCATCCTCCTTGCACCTTTGTGCCAA
W74664.RC        370 AGAAGCCAGGCCAGCAGCCCCGAGACCATCCTCCTTGNACCTTTGTGCCAA
AA443286.RC      303 AGAAGCCAGGCCAGCAGCCCCGAGACCATCCTCCTTGCACCTTTGTGCCAA
                  ++++++
<consen01>      736 AGAAGCCAGGCCAGCAGCCCCGAGACCATCCTCCTTGCACCTTTGTGCCAA
DNA47332.init    421 GAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAAAAA
W74664.RC        420 GAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAG
AA443286.RC      353 GAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAAAAA
                  ++++++
<consen01>      786 GAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAAGCAAAAAA

```

17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

FIGURE 3 C

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.